

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2005, 22:03:19 ; Search time 3103 Seconds
(without alignments)
2392.050 Million cell updates/sec

Title: US-10-672-108-5
Perfect score: 195
Sequence: 1 atgactactacaataat.....accatcatcactaagtgtga 195

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

- 1: gb_est1.*
- 2: gb_est2.*
- 3: gb_hcc.*
- 4: gb_est3.*
- 5: gb_est4.*
- 6: gb_est5.*
- 7: gb_est6.*
- 8: gb_gsl1.*
- 9: gb_gsl2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	39.8	20.4	764	8	BH947782	BH947782 obu80f07.
2	39	20.0	348	2	BE921962	BE921962 EST425731
3	39	20.0	925	7	CK249831	CK249831 EST733468
4	38.2	19.6	652	7	CF434974	CF434974 EST671319
5	38.2	19.6	685	7	CF474739	CF474739 EST683784
C 6	37.6	19.3	490	9	CR313060	CR313060 Medicago
C 7	37.6	19.3	763	5	BU255481	BU255481 603414348
C 8	37.4	19.2	407	6	CR279467	CR279467 qx45b06.y
9	37.4	19.2	486	6	CD192704	CD192704 MSI-0074T
C 10	37	19.0	270	4	BG931877	BG931877 hl22-167
11	37	19.0	525	6	CD084228	CD084228 MA3-9999U
12	36.6	18.8	592	8	AZ166486	AZ166486 SP 0088 A
C 13	36.6	18.8	869	9	AJ599051	AJ599051 Arabidops
C 14	36.4	18.7	888	9	CNS07C24	AL439814 T7 end of
C 15	36	18.5	411	6	CD534749	CD534749 40N18 Ara
C 16	36	18.5	436	5	BP627024	BP627024 BP627024
C 17	36	18.5	497	1	AV562914	AV562914 AV562914
18	36	18.5	520	6	CB281153	CB281153 ru35f11.y
19	36	18.5	529	8	AZ166362	AZ166362 SP 0088 A
C 20	36	18.5	534	6	CD533062	CD533062 29F11 Ara
21	36	18.5	547	1	AV553577	AV553577 AV553577
22	36	18.5	561	6	CB259450	CB259450 48-E9600-
C 23	36	18.5	576	6	CD530872	CD530872 08F11 Ara
24	36	18.5	616	9	CR310143	CR310143 Medicago

C 25	36	18.5	619	1	AI998882	AI998882 701547251
26	36	18.5	630	6	CD748424	CD748424 rw37b12.y
C 27	36	18.5	632	1	AV544982	AV544982 AV544982
28	36	18.5	812	8	BH240014	BH240014 EUTNA22TF
29	35.8	18.4	393	5	BQ112528	BQ112528 EST598104
C 30	35.8	18.4	445	6	CB825119	CB825119 rq32d02.y
31	35.6	18.3	381	7	CN038677	CN038677 nm_28.o17
C 32	35.4	18.2	624	5	BU420316	BU420316 603962412
33	35.2	18.1	588	2	AV998490	AV998490 AV998490
34	35.2	18.1	616	5	BW343719	BW343719 BW343719
35	35.2	18.1	630	5	BW052555	BW052555 BW052555
36	35.2	18.1	663	5	BW353709	BW353709 BW353709
37	35	17.9	360	1	AV198098	AV198098 AV198098
38	35	17.9	640	3	CNS0A0HV	CR294439 Arabidops
39	35	17.9	698	9	CR294439	CR294439 Medicago
40	35	17.9	814	8	BH684302	BH684302 BOMB235TF
C 41	34.8	17.8	653	4	BJ342138	BJ342138 BJ342138
42	34.8	17.8	801	8	BZ609984	BZ609984 WHABG35TR
43	34.8	17.8	888	7	CK273415	CK273415 EST719493
C 44	34.6	17.7	669	9	AG018250	AG018250 Homo sapi
C 45	34.6	17.7	680	9	CL636801	CL636801 CH243-8B1

ALIGNMENTS

RESULT 1
BH947782/c
LOCUS BH947782 764 bp DNA linear GSS 01-OCT-2002
DEFINITION obu80f07.g1 B.oleracea002 Brassica oleracea genomic, genomic survey
sequence.
ACCESSION BH947782
VERSION BH947782.1 GI:23427842
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 764)
AUTHORS Dalehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,
Nash,W., Rabinowicz,P.D. and Wilson,R.K.
TITLE Whole genome shotgun reads from Brassica oleracea
JOURNAL Unpublished (2002)
COMMENT Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Plate: obu80 row: f column: 07
Seq primer: -28RPpOT reverse
Class: shotgun
High quality sequence start: 88
High quality sequence stop: 551.

FEATURES
source
1..764
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/db_xref="taxon:3712"
/clone_lib="B.oleracea002"
/note="vector: pOTw13; Whole genome shotgun library from
flowering buds. DNA was purified from a crude nuclear
prep using Brassica oleracea T01000H3 buds provided by
Thomas Osborn at the University of Wisconsin. Genomic
DNA was provided by Pablo Rabinowicz (CSHL) and the
shotgun library prepared at Washington University Genome
Sequencing Center."

ORIGIN

Query Match 20.4%; Score 39.8; DB 8; Length 764;
Best Local Similarity 52.8%; Pred. No. 0.39;
Matches 86; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

OY 20 TCCTTAATGGTAAACATTGAAGCGAACAACACTACTGAAGCTGTTGATCTGCTACTG 79

Db 593 TACATAAATGATAAAATCTTGAGACTAGAGCCAACTGAAATTTTAAGTAATAAAACTA 534
Qy 80 CAGGGAAGCTTTCACACATAGCTTAACGACACCGTGTGACGGTGAATGACTTACG 139
Db 533 AAGAGAAACTATAAGAACAAAGGATTTGTAACAACGTCATGTGGTGGGTGGAACATC 474
Qy 140 ACATGCGCATAGACCTTTTACAGTTTACTGAACATCACCATCA 182
Db 473 ATGGTTTACTAGCAGCTCTATAGTAATAATCATAATTATTA 431

RESULT 2
BE921962 348 bp mRNA linear EST 07-MAR-2003
LOCUS EST425731 potato leaves and petioles Solanum tuberosum cDNA clone
DEFINITION CSTB17K20 5' sequence, mRNA sequence.
ACCESSION BE921962
VERSION BE921962.1 GI:10448038
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 348)
AUTHORS van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J.,
Unterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R.,
Ronning,C.M., Fry,W.E., Tankeley,S.D. and Baker,B.
TITLE Generation of ESTs from potato leaves and petioles
JOURNAL Unpublished (2000)
COMMENT Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/.

FEATURES
source
1..348
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTB17K20"
/tissue_type="leaflets and petioles"
/dev_stage="8 weeks old plants"
/lab_host="SOLR"
/clone_lib="potato leaves and petioles"
/note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
XhoI. Tissue was supplied by Dr. Fry (Cornell University).
Leaflets and petioles were isolated from 8 week old
greenhouse grown plants. The plants were watered and
fertilized freely. The tissue was immediately frozen in
liquid nitrogen."

ORIGIN
Query Match 20.0%; Score 39; DB 2; Length 348;
Best Local Similarity 50.8%; Pred. No. 0.6;
Matches 93; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

Qy 11 ACAATTAATCTTAATGGTAAACATTTGAAAGCGGAAACAACTACTCTGAAGCTGTTCATG 70
Db 124 AAAAAAATGCTGGAAGTCTCTGTTTGTGGAGATACAGTTACTGTGTGCGCGTCG 183
Qy 71 CTGCTACTGCGGAGAGTCTTCAACAATACGCTAACGACACGCGTGTGACGCGTAAT 130
Db 184 CCGCGCGCATGGTGAAGCGCGGAGCTGAAGAATGACGCGCGGTGTGGCGGTGGT 243
Qy 131 GGACTTACGAGATCGGACTTAAGACCTTTACAGTTACTGACATCACCATCATCACTAAG 190
Db 244 GTGGTTACAGAGGAGAGATGAAGATTTTTCAGTTGGAATCGTTGGCCACATGAGAAA 303

Qy 191 CTT 193
Db 304 CTT 306
RESULT 3
CK249831 925 bp mRNA linear EST 30-JUL-2004
LOCUS EST733468 potato callus cDNA library, normalized and full-length
DEFINITION Solanum tuberosum cDNA clone POCBB33 5' end, mRNA sequence.
ACCESSION CK249831
VERSION CK249831.1 GI:39801339
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 925)
AUTHORS Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
TITLE Generation of ESTs from potato callus tissue
JOURNAL Unpublished (2003)
COMMENT Other ESTs: EST733469
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/ .
Seq primer: AYT TAG GTG ACA CTA TAG.

FEATURES
source
1..925
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POCBB33"
/tissue_type="callus"
/lab_host="DH10B-Tona"
/clone_lib="potato callus cDNA library, normalized and
full-length"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Solanum tuberosum var.
Kennebec callus tissue grown on solid media."

ORIGIN
Query Match 20.0%; Score 39; DB 7; Length 925;
Best Local Similarity 50.8%; Pred. No. 0.71;
Matches 93; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

Qy 11 ACAATTAATCTTAATGGTAAACATTTGAAAGCGGAAACAACTACTCTGAAGCTGTTCATG 70
Db 170 AAAAAAATGCTGGAAGTCTCTGTTTGTGGAGATACAGTTACTGTGTGCGCGTCG 229
Qy 71 CTGCTACTGCGGAGAGTCTTCAACAATACGCTAACGACACGCGTGTGACGCGTAAT 130
Db 230 CCGCGCGCATGGTGAAGCGCGGAGCTGAAGAATGACGCGCGGTGTGGCGGTGGT 289
Qy 131 GGACTTACGAGATCGGACTTAAGACCTTTACAGTTACTGAACATCACCATCATCACTAAG 190
Db 290 GTGGTTCAGAGGAGAGATGAAGATTTTTCAGTTGGAATCGTTGGCCACATGAGAAA 349
Qy 191 CTT 193
Db 350 CTT 352

RESULT 4
CF434974 652 bp mRNA linear EST 04-SEP-2003
LOCUS EST671319 normalized cDNA library of onion Allium cepa cDNA clone
DEFINITION ACABP64, mRNA sequence.
ACCESSION CF434974

VERSION CF434974.1 GI:34457664
KEYWORDS EST.
SOURCE Allium cepa (onion)
ORGANISM Allium cepa

REFERENCE 1 (bases 1 to 652)
AUTHORS Havery, M.J., Cheung, P., Van Aken, S., Utterback, T. and Town, C.D.
TITLE Expressed Sequence Tags from a normalized library of mixed onion tissues (Allium cepa)

JOURNAL Unpublished (2003)
COMMENT Contact: Havery MJ
Department of Horticulture
USDA-ARS and University of Wisconsin
1575 Linden Drive, Madison, WI 53706, USA
Tel: 608-262-1830
Fax: 608-262-4743
Email: mjhavey@facstaff.wisc.edu
TIGR sequence name ACABP64TR. For more information:
http://haveylab.hort.wisc.edu

Seq primer: CAG GAA ACA GCT ATG ACC.
Location/Qualifiers
1..652
/organism="Allium cepa"
/mol_type="mRNA"
/cultivar="Red Creole(bulbs), unknown(callus), Ebano & Texas Legend(roots)"
/db_xref="taxon:4679"
/clone="ACABP64"
/tissue type="Callus, roots, and young bulbs"
/clone_lib="normalized cDNA library of onion"
/note="Vector: pCMVSPORT6.1-ccdb (Invitrogen); Site 1: EcoRV (5'); Site 2: NotI (3'); Equal molar amounts of mRNA from callus, roots, and young bulbs were combined to synthesize the library. Normalization to enrich for low-copy transcripts was performed by proprietary techniques of Invitrogen."

FEATURES
source

Query Match 19.6%; Score 38.2; DB 7; Length 652;
Best Local Similarity 51.5%; Pred. NO. 1.2;
Matches 88; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

Qy 24 TAATGGTAAACATTTGAAGCGGAACAACTACTGAAGCTGTTGATGCTGCTACTGCAGC 83
Db 400 TAATGGAAAGCCAAAGACAGCGAAGGTCCTAGCTGTTGATGCTGCTACTGCAGC 459
Qy 84 GAAAGCTCTTCAACAATACGCTTAACGCAACGCGTGTGACGCTGAATGGACTTACGAGA 143
Db 460 TAAAGTTAACTCAATATTAAGATGAAGATTCGACGACGAGATTCTGATGATGA 519
Qy 144 TCGGACTAAGACCTTTACAGTTACTGAACATCACATCATCCTAAAGCTTG 194
Db 520 TGAGGATATGTTGCGCTTGAAGTATGATGACGAGAGGATGAAGTAG 570

ORIGIN

Query Match 19.6%; Score 38.2; DB 7; Length 652;
Best Local Similarity 51.5%; Pred. NO. 1.2;
Matches 88; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

Qy 24 TAATGGTAAACATTTGAAGCGGAACAACTACTGAAGCTGTTGATGCTGCTACTGCAGC 83
Db 400 TAATGGAAAGCCAAAGACAGCGAAGGTCCTAGCTGTTGATGCTGCTACTGCAGC 459
Qy 84 GAAAGCTCTTCAACAATACGCTTAACGCAACGCGTGTGACGCTGAATGGACTTACGAGA 143
Db 460 TAAAGTTAACTCAATATTAAGATGAAGATTCGACGACGAGATTCTGATGATGA 519
Qy 144 TCGGACTAAGACCTTTACAGTTACTGAACATCACATCATCCTAAAGCTTG 194
Db 520 TGAGGATATGTTGCGCTTGAAGTATGATGACGAGAGGATGAAGTAG 570

RESULT 5

CF447439
LOCUS EST683784 normalized cDNA library of onion Allium cepa cDNA clone
DEFINITION ACAB11, mRNA sequence.

ACCESSION CF447439
VERSION CF447439.1 GI:34470141
KEYWORDS EST.

SOURCE Allium cepa (onion)
ORGANISM Allium cepa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Alliaceae; Allium.

REFERENCE 1 (bases 1 to 685)
AUTHORS Havery, M.J., Cheung, P., Van Aken, S., Utterback, T. and Town, C.D.
TITLE Expressed Sequence Tags from a normalized library of mixed onion

JOURNAL
COMMENT

tissues (Allium cepa)
Unpublished (2003)
Contact: Havery MJ
Department of Horticulture
USDA-ARS and University of Wisconsin
1575 Linden Drive, Madison, WI 53706, USA
Tel: 608-262-1830
Fax: 608-262-4743
Email: mjhavey@facstaff.wisc.edu
TIGR sequence name ACAB11TR. For more information:
http://haveylab.hort.wisc.edu
Seq primer: CAG GAA ACA GCT ATG ACC.
Location/Qualifiers
1..685
/organism="Allium cepa"
/mol_type="mRNA"
/cultivar="Red Creole(bulbs), unknown(callus), Ebano & Texas Legend(roots)"
/db_xref="taxon:4679"
/clone="ACAB11"
/tissue type="Callus, roots, and young bulbs"
/clone_lib="normalized cDNA library of onion"
/note="Vector: pCMVSPORT6.1-ccdb (Invitrogen); Site 1: EcoRV (5'); Site 2: NotI (3'); Equal molar amounts of mRNA from callus, roots, and young bulbs were combined to synthesize the library. Normalization to enrich for low-copy transcripts was performed by proprietary techniques of Invitrogen."

FEATURES
source

Query Match 19.6%; Score 38.2; DB 7; Length 685;
Best Local Similarity 51.5%; Pred. NO. 1.2;
Matches 88; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

ORIGIN

Qy 24 TAATGGTAAACATTTGAAGCGGAACAACTACTGAAGCTGTTGATGCTGCTACTGCAGC 83
Db 425 TAATGGAAAGCCAAAGACAGCGAAGGTCCTAGCTGTTGATGCTGCTACTGCAGC 484
Qy 84 GAAAGCTCTTCAACAATACGCTTAACGCAACGCGTGTGACGCTGAATGGACTTACGAGA 143
Db 485 TAAAGTTAACTCAATATTAAGATGAAGATTCGACGACGAGATTCTGATGATGA 544
Qy 144 TCGGACTAAGACCTTTACAGTTACTGAACATCACATCATCCTAAAGCTTG 194
Db 545 TGAGGATATGTTGCGCTTGAAGTATGATGACGAGAGGATGAAGTAG 595

RESULT 6

CR313060/c
LOCUS CR313060.1 GI:44859204
DEFINITION Medicago truncatula BAC ends cultivar Jemalong A17 of Medicago truncatula, genomic survey sequence.

ACCESSION CR313060
VERSION CR313060.1 GI:44859204
KEYWORDS GSS.

SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

REFERENCE 1 (bases 1 to 490)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (25-FEB-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

FEATURES
source

1..490
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/cultivar="Jemalong A17"
/db_xref="taxon:3880"

```
/clone lib="MTE1"
/notes=Vector; pIndigoBAC ; Site 1: EcoRI ; Site 2: EcoRI
; Debelle F. and Chalhoub B.-Genoscope sequence ID :
mte1-35P4PM1"
```

```
ORIGIN
Query Match      19.3%; Score 37.6; DB 9; Length 490;
Best Local Similarity 50.6%; Pred. No. 1.7;
Matches 91; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

Qy 7 ACTTACAAATTAATCCTTAATGGTAAACATTTGAAGGCGAACAACACTACTGAAGCTGTT 66
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 442 ACTCACAATTTGTAACAAAAAATAAAAAAACAAGAGATTTAAAAAAGAAGCATTTG 383
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 67 GATGCTGCTACTGCAGCGAAGTCTTCAACAATACGCTAACACACACGCGTGTGGCGT 126
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 382 ATTAGTGGCAATCTTGTGTAATACATAAAATAAAAGAACCGAATGAAGAATATACT 323
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 127 GAATGGACTTACGACGATGCGACTTAAGACCTTTACAGTTACTGAACATCACCATCATCAC 186
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 322 AATAAATGATTCACGTGCGAAGAAACAAACATAAACTCATCATCATCATCATCAC 263
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 7
BU255481/c
LOCUS      763 bp mRNA linear EST 26-NOV-2002
DEFINITION 603414348F1 CSEQCHN38 Gallus gallus CDNA clone CHEST336p12 5', mRNA
sequence.
ACCESSION BU255481
VERSION   BU255481.1 GI:25512682
KEYWORDS  EST.
SOURCE    Gallus gallus (chicken)
ORGANISM  Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 763)
AUTHORS   Boardman,P.E., Sanz-Exquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
A Comprehensive Collection of Chicken cDNAs
TITLE     Curr. Biol. 12 (22), 1965-1969 (2002)
JOURNAL   22335534
MEDLINE   12445392
PUBMED    12445392
COMMENT   Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
```

```
FEATURES
source
1. .763
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="CHEST336p12"
/dev_stage="36"
/lab_host="DH10B"
/clone_lib="CSEQCHN38"
/notes="Organ: limbs; Vector: pBluescript II KS(+); Site_1:
EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunt-ended, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
```

ORIGIN

```
Query Match      19.3%; Score 37.6; DB 5; Length 763;
Best Local Similarity 54.3%; Pred. No. 1.8;
Matches 76; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

Qy 38 TGAAGCGGAAACAACACTACTGAAGCTGTGTCATGCTCTACTGCAGCGAAGTCTTCAAAC 97
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 417 TGTGACCTTGAGCGACTCTTGGTGGTGTACGAGGCTGTTCCAGTGTGAGATTGCCAC 358
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 98 AATACGCTAACGACAAACGCGTGTGACGGTGAATGGACTTTACGACGATGCGACTTAACACCT 157
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 357 ATCCAGATAACCAACAATCCAGAGGCTGAATGGACACAAAGACGATCTTTCAAGTTCCC 298
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 158 TTACACTTTACTGAACATCAC 177
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 297 TTCCAATCCCTGACATTCTC 278
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 8 CB279467/c

```
LOCUS      407 bp mRNA linear EST 26-FEB-2003
DEFINITION rq45D06.y1 Heterodera glycines J3 Heterodera glycines cDNA 5'
similar to TR:O02246 O02246 ZK270.1 PROTEIN. [1] ; mRNA sequence.
ACCESSION CB279467
VERSION   CB279467.1 GI:28567844
KEYWORDS  EST.
SOURCE    Heterodera glycines
ORGANISM  Heterodera glycines
Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
Tylenchoidea; Heteroderidae; Heteroderinae; Heterodera.
REFERENCE 1 (bases 1 to 407)
AUTHORS   McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J.,
Wyllie,T., Dante,M., Marra,M., Hillier,L., Bennett,J., Franklin,C.,
Tsagarisshilli,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T.,
Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
The Washington Univ. Nematode EST Project, 1999
Unpublished (1999)
Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This library was generated by cloning cDNAs directionally into
Uni-ZAP(Stratagene) (T3 primer/EcoRI are at the 5'-end and T7/XhoI
are at the 3'-end). The library was excised in pBluescript
SK(+) and normalized (Bonaldo et al 1996 Genome Research
6:791-806). Library constructed by Thomas Baum (tbaum@tastate.edu),
Iowa State University, Plant Pathology Department and Jeff
McDermott (jpmcdex@iastate.edu).
```

FEATURES

```
source
1. .407
/organism="Heterodera glycines"
/mol_type="mRNA"
/db_xref="taxon:51029"
/sex="mixed"
/tissue_type="whole organism"
/dev_stage="3rd stage juvenile"
/lab_host="DH10B"
/clone_lib="Heterodera glycines J3"
/notes="Vector: pBluescript SK+ (Stratagene); Site 1: XhoI;
Site2: EcoRI; This library was generated by cloning cDNAs
directionally into Uni-ZAP(Stratagene) (T3 primer/EcoRI
```


Db	111	ACGATCAGCAGCAGCATGATGAAGAAGAGGAGGAGGTGATGCTGTCAGAACAAAGATGATG	52
Qy	161	CAGTTACTGACATCACCATCACTAA	189
Db	51	AAGATAATGAACAACAACAACAACA	23
RESULT 11			
LOCUS	CD084228		
DEFINITION	MA3-9999U-V279-B05-U-B MA3-0001 Schistosoma mansoni cDNA clone		
ACCESSION	CD084228		
VERSION	CD084228.1	GI:34635208	
KEYWORDS	EST.		
SOURCE	Schistosoma mansoni		
ORGANISM	Schistosoma mansoni		
REFERENCE	Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.		
AUTHORS	1 (bases 1 to 525) Verjovski-Almeida, S., DeMarco, R., Martins, E.A.L., Guimaraes, P.E.M., Ojopi, E.P.B., Paquola, A.C.M., Piazza, J.P., Nishiyama, M.Y. Jr., Kitajima, J.P., Adamson, R.E., Ashton, P.D., Bonaldo, M.F., Coulson, P.S., Dillon, G.P., Farias, L.P., Gregorio, S.P., Ho, P.L., Leite, R.A., Malaquias, L.C.C., Marques, R.C.P., Miyasato, P.A., Nascimento, A.L.T.O., Ohlweiler, F.P., Reis, E.M., Ribeiro, M.A., Sa, R.G., Stukart, G.C., Soares, M.B., Gargioni, C., Kawano, T., Rodrigues, V., Madeira, A.M.B.N., Wilson, R.A., Menck, C.F.M., Setubal, J.C., Leite, L.C.C. and Dias-Neto, E.		
TITLE	Transcriptome analysis of the acelomate human parasite Schistosoma mansoni		
JOURNAL	Nat. Genet. 35 (2), 148-157 (2003)		
MEDLINE	22879926		
PUBMED	12973350		
COMMENT	Contact: Dr. Sergio Verjovski-Almeida Departamento de Bioquimica Instituto de Quimica - Universidade de Sao Paulo Av. Prof. Lineu Prestes 748 sala 1200, 05508-900 Sao Paulo - SP, Brasil Tel: +55-11-3091-2173 Fax: +55-11-3091-2186 Email: verjo@iq.usp.br This sequence was derived from the FAPESP Schistosoma mansoni EST Genome Project. All sequences in the project were assembled and annotated. This entry and all the assembled sequences can be seen in the following URL http://bioinfo.iq.usp.br/schisto/ Plate: MA3-9999U-V279 row: 5 column: B.		
FEATURES	1..525		
source	Location/Qualifiers		
ORIGIN	1..525		
Query Match	19.0%;	Score 37;	DB 6; Length 525;
Best Local Similarity	53.0%;	Pred. No. 2.6;	
Matches	79;	Conservative	0; Mismatches 70; Indels 0; Gaps 0;
Qy	41	AAGCGCAACAACTACTGAAGCTGTTGATGCTGCTACTGCGAGCAAGTCTTCAACAAT	100
Db	163	AAGTAGGACGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	222
Qy	101	ACGCTAACGCAACGGTGTTCACGTGAATGCTTACGACGCTTACGACGCTTACGACCTTTA	160
Db	223	ACGATGACGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	282
Qy	161	CAGTTACTGAACATCACCATCACTAA	189
Db	111	ACGATCAGCAGCAGCATGATGAAGAAGAGGAGGAGGTGATGCTGTCAGAACAAAGATGATG	52
Qy	161	CAGTTACTGACATCACCATCACTAA	189
Db	51	AAGATAATGAACAACAACAACAACAACA	23
RESULT 12			
LOCUS	AZ166486		
DEFINITION	SP_0088_A2_C11_T7A Strongylocentrotus purpuratus, purple sea urchin, sperm genomic BAC library Strongylocentrotus purpuratus genomic clone Plate=88 Col=22 Row=E, genomic survey sequence.		
ACCESSION	AZ166486		
VERSION	AZ166486.1	GI:8336815	
KEYWORDS	GSS.		
SOURCE	Strongylocentrotus purpuratus		
ORGANISM	Strongylocentrotus purpuratus		
REFERENCE	Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa; Echinoidea; Euechinoidea; Echinacea; Echinoidea; Strongylocentrotidae; Strongylocentrotus.		
AUTHORS	1 (bases 1 to 592) Cameron, R.A., Mahairas, G., Rast, J.P., Martinez, P., Biondi, T.R., Swartzell, S., Wallace, J.C., Poustka, A.J., Livingston, B.T., Wray, G.A., Ettensohn, C.A., Lehrach, H., Britten, R.J., Davidson, E.H. and Hood, L.		
TITLE	A sea urchin genome project: Sequence scan, virtual map, and additional resources		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (17), 9514-9518 (2000)		
MEDLINE	20402566		
PUBMED	10920195		
COMMENT	Contact: Cameron, RA, Davidson, EH, Hood, L Division of Biology 156-29 California Institute of Technology Pasadena California 91125, USA Tel: (626) 395-8421 Fax: (626) 793-3047 Email: acameron@caltech.edu Plate: 88 row: E column: 22 Seq primer: T7 Class: BAC ends High quality sequence stop: 592.		
FEATURES	1..592		
source	Location/Qualifiers		
ORIGIN	1..592		
Query Match	18.8%;	Score 36.6;	DB 8; Length 592;
Best Local Similarity	53.6%;	Pred. No. 3.5;	
Matches	75;	Conservative	0; Mismatches 65; Indels 0; Gaps 0;
Qy	49	ACAACACTGAAGCTGTGTGATGCTGCTACTGCGAGCAAGTCTTCAACAATAGCTAAC	108
Db	359	ACTACTACCAATACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTAC	418
Qy	109	GACACGGTGTGACCGTGAATGACCTTACGACGATGCGACTTAAGCCTTTACAGTTACT	168
Db	419	TACTACAGNACAAATGATCCATCTAATTATGCTCTATTACTATAAATACTACTACT	478
Qy	169	GAACATCACCATCATCACTA	188
Db	479	ACCTAATACTACTACTACTA	498
RESULT 13			
LOCUS	AJ599051/c		
DEFINITION	Arabidopsis thaliana T-DNA flanking sequence, right border, clone 479E08, genomic survey sequence.		
ACCESSION	AJ599051		

```

VERSION AJ599051.1 GI:37948679
KEYWORDS GSS; right border; T-DNA flanking sequence.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE
AUTHORS Brunaud, V., Balzerque, S., Dubreucq, B., Aubourg, S., Samson, F.,
Chauvin, S., Bechtold, N., Cruaud, C., DeRose, R., Pelletier, G.,
Lepiniec, L., Caboche, M., and Lecharny, A.
TITLE T-DNA integration into the Arabidopsis genome depends on sequences
of pre-insertion sites
JOURNAL EMBO Rep. 3 (12), 1152-1157 (2002)
MEDLINE 22363535
PUBMED 12446565
REFERENCE
AUTHORS 2 (bases 1 to 869)
Balzerque, S.
TITLE Direct Submission
JOURNAL Submitted (23-OCT-2003) Balzerque S., UMRGV, INRA/CNRS, 2 rue
Gaston Cremieux, 91057 Evry cedex, FRANCE
COMMENT PCR was performed on DNA from transformants of Arabidopsis thaliana
plants from INRA (Versailles). The DNA fragment (e) resulting from
the PCR were directly sequenced from the left or the right border
to determine the genomic sequence flanking the insertion. T-DNA
derived sequences were removed. Information to order the
corresponding mutant line and a link to a database providing a
graphical display of the insertion site are available at
http://dbsgap.versailles.inra.fr/publiclines/. This sequence has
been generated in the framework of the French plant genomics
program 'Genoplante' (http://www.genoplante.com and
http://genoplante-info.infobiogen.fr).
FEATURES
source
1..869
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/cultivar="Wassillewskija"
/db_xref="taxon:3702"
/clone="479E08"
/clone_lib="Arabidopsis thaliana T-DNA insertion lines"
misc_feature
1..869
/note="T-DNA flanking sequence
right border"
ORIGIN
Query Match 18.8%; Score 36.6; DB 9; Length 869;
Best Local Similarity 57.4%; Pred. No. 3.8;
Matches 66; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 51 AACTACTGAAGCTGTTGATGCTGCTACGCGAAGACTCTTCAACAATACGCTAACGA 110
DB 645 ATCTCTTGAATCGGAAGAAGCTTCTCCGTCGCGAAGCAATCGAAGAAGCTTACGC 586
QY 111 CAACGGTGTGACGGTGAATGACCTTACGACGATCGGACTTAAGACTTTTACAGTT 165
DB 585 GTGCTTCTGCCACCGTATTTCGGTGATGACGCGGATCGAGATCTTTAAAGCT 531
RESULT 14
CNS07C24 888 bp DNA linear GSS 08-JUL-2001
LOCUS T7 end of clone BD0AA008B01 of library BD0AA from strain CBS 94 of
DEFINITION Candida tropicalis, genomic survey sequence.
ACCESSION AL439814
VERSION AL439814.1 GI:12223226
KEYWORDS GSS.
SOURCE Candida tropicalis
ORGANISM Candida tropicalis
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE
AUTHORS 1 (bases 1 to 888)
Souciet, J.L., Aigle, M., Artiguenave, F., Blandin, G.,
Bolotin-Fukuhara, M., Bon, E., Brottier, P., Casaregola, S.,
de-Montigny, J., Dujon, B., Durrens, P., Lepingle, A., Llorente, B.,
Malpertuy, A., Neuveglise, C., Ozier-Kalogeropoulos, O., Potier, S.,
Saurin, W., Tekala, P., Toffano-Nioche, C., Wesolowski-Louvel, M.,
Wincker, P. and Weissenbach, J.
TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of
yeast species for molecular evolution studies
JOURNAL FEBS Lett. 487 (1), 3-12 (2000)
MEDLINE 20584711
PUBMED 11152876
REFERENCE
AUTHORS 2 (bases 1 to 888)
Blandin, G., Ozier-Kalogeropoulos, O., Wincker, P., Artiguenave, F. and
Dujon, B.
TITLE Genomic exploration of the hemiascomycetous yeasts: 16. Candida
tropicalis
JOURNAL FEBS Lett. 487 (1), 91-94 (2000)
MEDLINE 20584726
PUBMED 11152891
REFERENCE
AUTHORS 3 (bases 1 to 888)
Genoscope.
TITLE Direct Submission
JOURNAL Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
2 rue Gaston Cremieux, CP 5706, 91057 Evry cedex, FRANCE. (E-mail :
segrif@genoscope.cns.fr - Web :
http://www.genoscope.cns.fr)
COMMENT This GSS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazii, Zygosaccharomyces rouxii,
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of
the other extremity of this insert.
FEATURES
source
1..888
/organism="Candida tropicalis"
/mol_type="genomic DNA"
/strain="CBS 94"
/db_xref="taxon:5482"
/clone="BD0AA008B01"
/clone_lib="BD0AA"
/note="end : T7"
ORIGIN
Query Match 18.7%; Score 36.4; DB 9; Length 888;
Best Local Similarity 56.8%; Pred. No. 4.3;
Matches 67; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
QY 46 GAAACAACTACTGAAGCTGTTGATGCTGCTACTGTCAGCGAAGTCTTCAACAATACGCT 105
DB 650 GAAGAAACCCAGCGAAGCTGCTGTAACCAACCACTGCTGCTGAAGAAACCCAGCGAAGCCGCT 591
QY 106 AACGACAAACGGTGTGACGGTGAATGACCTTACGACGATCGGACTTACGACCTTTACAG 163
DB 590 GAAACCAACCACTGCTGCTGTAAGAAACCACTGAAGCCGCTGAACCAACCACTGCTGCTG 533
RESULT 15
CNS34749/c 411 bp mRNA linear EST 31-DEC-2003
LOCUS 40N18 Arabidopsis Leaf Senescence Library Arabidopsis thaliana CDNA
DEFINITION 3', mRNA sequence.
ACCESSION CD534749
VERSION CD534749.1 GI:40454761
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE
AUTHORS 1 (bases 1 to 411)
Guo, Y., Cai, Z. and Gan, S.
TITLE Transcriptome of Arabidopsis leaf senescence

```

JOURNAL
COMMENT

Plant Cell Environ. 27 (5), 521-549 (2004)
Contact: Susheng Gan
Department of Horticulture
Cornell University
119 Plant Science, Cornell University, Ithaca, NY 14853-5904, USA
Tel: 607 254 5418
Fax: 607 255 0599
Email: sg288@cornell.edu
Insert Length: 411 Std Error: 0.00
Seq primer: T7
POLYA=Yes.

FEATURES

source Location/Qualifiers
1..411
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Landsberg erecta"
/db_xref="taxon:3702"
/tissue_type="Leaf"
/dev_stage="Yellow Leaf With Greenish Base Area"
/lab_host="E. coli"
/clone_lib="Arabidopsis Leaf Senescence Library"
/note="Organ: Rosette Leaf; Vector: pBluscript SKII+;
Site 1: EcoRI; Site 2: EcoRI; Senescent rosette leaves #5
and #6 (counted from the bottom) were harvested and
immediately frozen in liquid N2. The leaves were visibly
yellow excepted for the leaf base areas that were still
greenish."

ORIGIN

Query Match 18.5%; Score 36; DB 6; Length 411;
Best Local Similarity 58.3%; Pred. No. 5;
Matches 63; Conservative 0; Mismatches 45; Indels 0; Gaps 0;
Qy 52 ACTACTGAAGCTGTTGATGCTGCTACTGCGAGCGAAGCTTCAACAAATACGCTAAGCAC 111
Db 394 ACTCTTAAATCTGACGATGCAACCCGTCGCGAAACTTATCGAAGAAAGGCTTATGGT 335
Qy 112 AACGGTGTTCACGGTGAATGGACTTACGACGATGCGACTAAGACCTTT 159
Db 334 GTTGCTTGGATGCTGTGTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 287

Search completed: August 10, 2005, 23:51:00
Job time : 3106 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2005, 13:36:38 ; Search time 430 Seconds
(without alignments)

2684.534 Million cell updates/sec

Title: US-10-672-108-5

Perfect score: 195

Sequence: 1 atgactactacaataat.....accatcatcactaagtgtga 195

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_16Dec04:*

- 1: geneseqn1980s:*
- 2: geneseqn1980s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004as:*
- 13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	195	100.0	195	4 AAC86821	AAC86821 Nucleotid
2	191.8	98.4	195	4 AAC86827	AAC86827 Nucleotid
3	191.8	98.4	195	4 AAC86820	AAC86820 Nucleotid
4	190.2	97.5	195	4 AAC86828	AAC86828 Nucleotid
5	190.2	97.5	195	4 AAC86829	AAC86829 Nucleotid
6	188.6	96.7	195	4 AAC86825	AAC86825 Nucleotid
7	187	95.9	195	4 AAC86822	AAC86822 Nucleotid
8	187	95.9	195	4 AAC86823	AAC86823 Nucleotid
9	187	95.9	195	4 AAC86824	AAC86824 Nucleotid
10	183.8	94.3	195	4 AAC86826	AAC86826 Nucleotid
11	163.8	84.0	1308	2 AAQ50453	AAQ50453 Sequence
12	162.6	83.4	396	1 AAN94677	AAN94677 Protein G
13	162.6	83.4	396	1 AAN94677	AAN94677 Protein G
14	162.6	83.4	396	2 AAQ64749	AAQ64749 IGG-bind
15	162.6	83.4	435	1 AAN91100	AAN91100 Protein G
16	162.6	83.4	435	2 AAQ06008	AAQ06008 Sequence
17	162.6	83.4	435	2 AAQ10005	AAQ10005 Type 3 GX
18	162.6	83.4	435	2 AAQ64647	AAQ64647 IGG-bind
19	162.6	83.4	447	1 AAN91103	AAN91103 Protein G
20	162.6	83.4	447	2 AAQ06006	AAQ06006 Sequence

21	162.6	83.4	447	2 AAQ10004	AAQ10004 Type 1 St
22	162.6	83.4	447	2 AAQ64645	AAQ64645 IGG-bind
23	162.6	83.4	495	1 AAN91108	AAN91108 Protein G
24	162.6	83.4	495	2 AAQ06016	AAQ06016 Sequence
25	162.6	83.4	495	2 AAQ10011	AAQ10011 Type 9 GX
26	162.6	83.4	495	2 AAQ64748	AAQ64748 IGG-bind
27	162.6	83.4	600	1 AAN91104	AAN91104 Protein G
28	162.6	83.4	600	2 AAQ06007	AAQ06007 Sequence
29	162.6	83.4	600	2 AAQ10003	AAQ10003 Type 2 GX
30	162.6	83.4	600	2 AAQ64646	AAQ64646 IGG-bind
31	162.6	83.4	606	2 AAQ10009	AAQ10009 Type 8 GX
32	162.6	83.4	750	2 AAQ10008	AAQ10008 Type 6 GX
33	162.6	83.4	798	1 AAN91099	AAN91099 Protein G
34	162.6	83.4	798	2 AAQ06009	AAQ06009 Sequence
35	162.6	83.4	798	2 AAQ10007	AAQ10007 Type 4 GX
36	162.6	83.4	798	2 AAQ64648	AAQ64648 IGG-bind
37	162.6	83.4	810	1 AAN94674	AAN94674 Protein G
38	162.6	83.4	810	2 AAQ06011	AAQ06011 Sequence
39	162.6	83.4	810	2 AAQ64649	AAQ64649 IGG-bind
40	162.6	83.4	1469	1 AAN70757	AAN70757 Sequence
41	162.6	83.4	1949	2 AAQ06018	AAQ06018 Sequence
42	162.6	83.4	1950	1 AAN91093	AAN91093 Protein G
43	162.6	83.4	1950	1 AAN91094	AAN91094 Protein G
44	162.6	83.4	1950	2 AAQ06017	AAQ06017 Sequence
45	162.6	83.4	1950	2 AAQ10001	AAQ10001 Sequence

ALIGNMENTS

RESULT 1

AAC86821
ID AAC86821 standard; DNA; 195 BP.

XX AAC86821;

XX 02-APR-2001 (first entry)

XX Nucleotide sequence of a B1 domain of protein G with mutation E27A.

XX B1 domain; protein G; Fab fragment; immunoglobulin G; IGG; Fc fragment;
XX knobs-into-holes binding site; ss.

XX Synthetic.

XX Streptococcus sp.

XX Key Location/Qualifiers

FT CDS 1..189

FT /*tag= a

FT /product= "B1 domain of protein G"

FT mat_peptide 4..189

FT /*tag= c

XX WO200074728-A1.

XX 14-DEC-2000.

XX 02-JUN-2000; 2000WO-US015366.

XX 04-JUN-1999; 99US-00326342.

XX (UYDU-) UNIV DUKE.

XX Hellinga HW, Sloan DJ;

XX WPI; 2001-049994/06.

XX P-PSDB; AAB30886.

XX B1 domain polypeptide of bacterial protein G, useful for detection and
XX purification of antibody fragments, exhibits binding activity for Fab
XX fragment of an IGG but does not bind Fc fragment.

XX Claim 34; Page 73; 83pp; English.

XX The present sequence encodes a mutated B1 domain of the bacterial protein
CC G. The B1 domain exhibits binding activity for a Fab fragment of an
CC immunoglobulin (Ig) G but exhibits no binding activity for a Fc fragment
CC of an IgG. The B1 domain may further comprise a disrupted or mutated
CC "knobs-into-holes" binding site for a Fc fragment of an IgG. The Glu27
CC residue of the mature B1 domain acts as a charged knob on the surface of
CC the B1 domain, inserting a polar hole on the Fc fragment. A mutation of
CC this residue virtually abolishes stable complex formation. The B1 domain
CC can be immobilized to a solid phase support, and be used for purifying Fc
CC and Fab fragments of IgG from a warm-blooded vertebrate by affinity
CC chromatography. The B1 domain is also useful for detecting IgG or its
CC fragments in a clinical sample
XX
SQ Sequence 195 BP; 68 A; 42 C; 37 G; 48 T; 0 U; 0 Other;

Query Match 100.0%; Score 195; DB 4; Length 195;
Best Local Similarity 100.0%; Pred. No. 9e-47;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACTACTTACAAATTAATCCCTTAATGGTAAACATTGAAAGCGGAAACAACTACTGAA 60
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1 ATGACTACTTACAAATTAATCCCTTAATGGTAAACATTGAAAGCGGAAACAACTACTGAA 60
QY 61 GCTGTTGATGCTGCTACTGCGAGCGAAGTCTTCAAAACAATACGCTAACGACACGGTGTT 120
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
61 GCTGTTGATGCTGCTACTGCGAGCGAAGTCTTCAAAACAATACGCTAACGACACGGTGTT 120
QY 121 GACGGTGAATGGACTTACGACGATGCGACTAAGACCTTTACGTTACTGAACATCACCAT 180
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121 GACGGTGAATGGACTTACGACGATGCGACTAAGACCTTTACGTTACTGAACATCACCAT 180
QY 181 CATCACTAAGCTTGA 195
DB ||||||||||||||||||||
181 CATCACTAAGCTTGA 195

RESULT 2
AAC86827
ID AAC86827 standard; DNA; 195 BP.
XX
AC AAC86827;
XX
DT 02-APR-2001 (first entry)
XX
DE Nucleotide sequence of a mutated B1 domain of protein G.
XX
KW B1 domain; protein G; Fab fragment; immunoglobulin G; IgG; Fc fragment;
KW knobs-into-holes binding site; ss.
XX
OS Synthetic.
OS Streptococcus sp.

XX Key Location/Qualifiers
FH CDS 1..189
FT /*tag= a
FT /product= "B1 domain of protein G"
FT mat_peptide 4..189
FT /*tag= c

XX WO200074728-A1.
XX
XX 14-DEC-2000.
XX
XX 02-JUN-2000; 2000WO-US015366.
XX
XX 04-JUN-1999; 99US-00326342.
XX
XX (UYDU-) UNIV DUKE.
XX
XX Hellinga HW, Sloan DJ;
XX
XX WPI; 2001-049994/06.

DR P-PSDB; AAB30894.
XX
XX B1 domain polypeptide of bacterial protein G, useful for detection and
PT purification of antibody fragments, exhibits binding activity for Fab
PT fragment of an IgG but does not bind Fc fragment.
XX
PS Claim 34; Page 78-79; 83pp; English.

XX The present sequence encodes a mutated B1 domain of the bacterial protein
CC G. The B1 domain exhibits binding activity for a Fab fragment of an
CC immunoglobulin (Ig) G but exhibits no binding activity for a Fc fragment
CC of an IgG. The B1 domain may further comprise a disrupted or mutated
CC "knobs-into-holes" binding site for a Fc fragment of an IgG. The Glu27
CC residue of the mature B1 domain acts as a charged knob on the surface of
CC the B1 domain, inserting a polar hole on the Fc fragment. A mutation of
CC this residue virtually abolishes stable complex formation. The B1 domain
CC can be immobilized to a solid phase support, and be used for purifying Fc
CC and Fab fragments of IgG from a warm-blooded vertebrate by affinity
CC chromatography. The B1 domain is also useful for detecting IgG or its
CC fragments in a clinical sample

SQ Sequence 195 BP; 68 A; 41 C; 36 G; 50 T; 0 U; 0 Other;

Query Match 98.4%; Score 191.8; DB 4; Length 195;
Best Local Similarity 99.0%; Pred. No. 7.7e-46;
Matches 193; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGACTACTTACAAATTAATCCCTTAATGGTAAACATTGAAAGCGGAAACAACTACTGAA 60
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1 ATGACTACTTACAAATTAATCCCTTAATGGTAAACATTGAAAGCGGAAACAACTACTGAA 60
QY 61 GCTGTTGATGCTGCTACTGCGAGCGAAGTCTTCAAAACAATACGCTAACGACACGGTGTT 120
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
61 GCTGTTGATGCTGCTACTGCGAGTAAAGTCTTCAAAACAATACGCTAACGACACGGTGTT 120
QY 121 GACGGTGAATGGACTTACGACGATGCGACTAAGACCTTTACGTTACTGAACATCACCAT 180
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121 GACGGTGAATGGACTTACGACGATGCGACTAAGACCTTTACGTTACTGAACATCACCAT 180
QY 181 CATCACTAAGCTTGA 195
DB ||||||||||||||||||||
181 CATCACTAAGCTTGA 195

RESULT 3
AAC86820
ID AAC86820 standard; DNA; 195 BP.
XX
AC AAC86820;
XX
DT 02-APR-2001 (first entry)
XX
DE Nucleotide sequence of the native B1 domain of protein G.
XX
KW B1 domain; protein G; Fab fragment; immunoglobulin G; IgG; Fc fragment;
KW knobs-into-holes binding site; ss.
XX
OS Streptococcus sp.

XX Key Location/Qualifiers
FH CDS 1..189
FT /*tag= a
FT /product= "B1 domain of protein G"
FT mat_peptide 4..189
FT /*tag= c

XX WO200074728-A1.
XX
XX 14-DEC-2000.
XX
XX 02-JUN-2000; 2000WO-US015366.
XX
XX 04-JUN-1999; 99US-00326342.

```

XX PA (UYDU-) UNIV DUKE.
XX PI Hellinga HW, Sloan DJ;
XX DR WPI; 2001-049994/06.
XX DR P-PSDB; AAB30883.
XX PT B1 domain polypeptide of bacterial protein G, useful for detection and
XX PT purification of antibody fragments, exhibits binding activity for Fab
XX PT fragment of an IgG but does not bind Fc fragment.
XX PS Disclosure; Page 50; 83pp; English.
XX CC The present sequence encodes a native B1 domain of the bacterial protein
XX CC G. The B1 domain exhibits binding activity for a Fab fragment of an
XX CC immunoglobulin (Ig) G but exhibits no binding activity for a Fc fragment
XX CC of an IgG. The B1 domain may further comprise a disrupted or mutated
XX CC "knobs-into-holes" binding site for a Fc fragment of an IgG. The Glu27
XX CC residue of the mature B1 domain acts as a charged knob on the surface of
XX CC the B1 domain, inserting a polar hole on the Fc fragment. A mutation of
XX CC this residue virtually abolishes stable complex formation. The B1 domain
XX CC can be immobilized to a solid phase support, and be used for purifying Fc
XX CC and Fab fragments of IgG from a warm-blooded vertebrate by affinity
XX CC chromatography. The B1 domain is also useful for detecting IgG or its
XX CC fragments in a clinical sample
XX SQ Sequence 195 BP; 70 A; 41 C; 36 G; 48 T; 0 U; 0 Other;

Query Match 98.4%; Score 191.8; DB 4; Length 195;
Best Local Similarity 99.0%; Pred. No. 7.7e-46;
Matches 193; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGACTACTTACAAATTAATCTTAATGGTTAAACATTGAAGCGGAACAACTACTGAA 60
Db 1 ATGACTACTTACAAATTAATCTTAATGGTTAAACATTGAAGCGGAACAACTACTGAA 60
Qy 61 GCTGTTGATGCTGCTACTGCGAGCAAAAGTCTTCAACAATACGCTAACGCAACGGTGT 120
Db 61 GCTGTTGATGCTGCTACTGCGAGCAAAAGTCTTCAACAATACGCTAACGCAACGGTGT 120
Qy 121 GACGGTGAATGAGCTTACGACGATGCGACTAAGACCTTTACAGTTACTGAAACATCACCAT 180
Db 121 GACGGTGAATGAGCTTACGACGATGCGACTAAGACCTTTACAGTTACTGAAACATCACCAT 180
Qy 181 CATCACTAAGCTTGA 195
Db 181 CATCACTAAGCTTGA 195

RESULT 4
AAC86828
ID AAC86828 standard; DNA; 195 BP.
XX AC AAC86828;
XX DT 02-APR-2001 (first entry)
XX DE Nucleotide sequence of a mutated B1 domain of protein G.
XX KW B1 domain; protein G; Fab fragment; immunoglobulin G; IgG; Fc fragment;
XX KW knobs-into-holes binding site; ss.
XX OS Synthetic.
XX OS Streptococcus sp.
XX FH Key Location/Qualifiers
XX FT CDS 1..189
XX FT /*tag= a
XX FT /product= "B1 domain of protein G"
XX FT mat_peptide 4..189
XX FT /*tag= c
XX

```

```

PN WO200074728-A1.
XX PD 14-DEC-2000.
XX PF 02-JUN-2000; 2000WO-US015366.
XX PR 04-JUN-1999; 99US-00326342.
XX PA (UYDU-) UNIV DUKE.
XX PI Hellinga HW, Sloan DJ;
XX DR WPI; 2001-049994/06.
XX DR P-PSDB; AAB30925.
XX PT B1 domain polypeptide of bacterial protein G, useful for detection and
XX PT purification of antibody fragments, exhibits binding activity for Fab
XX PT fragment of an IgG but does not bind Fc fragment.
XX PS Claim 34; Page 79; 83pp; English.
XX CC The present sequence encodes a mutated B1 domain of the bacterial protein
XX CC G. The B1 domain exhibits binding activity for a Fab fragment of an
XX CC immunoglobulin (Ig) G but exhibits no binding activity for a Fc fragment
XX CC of an IgG. The B1 domain may further comprise a disrupted or mutated
XX CC "knobs-into-holes" binding site for a Fc fragment of an IgG. The Glu27
XX CC residue of the mature B1 domain acts as a charged knob on the surface of
XX CC the B1 domain, inserting a polar hole on the Fc fragment. A mutation of
XX CC this residue virtually abolishes stable complex formation. The B1 domain
XX CC can be immobilized to a solid phase support, and be used for purifying Fc
XX CC and Fab fragments of IgG from a warm-blooded vertebrate by affinity
XX CC chromatography. The B1 domain is also useful for detecting IgG or its
XX CC fragments in a clinical sample
XX SQ Sequence 195 BP; 69 A; 41 C; 35 G; 50 T; 0 U; 0 Other;

Query Match 97.5%; Score 190.2; DB 4; Length 195;
Best Local Similarity 98.5%; Pred. No. 2.2e-45;
Matches 192; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGACTACTTACAAATTAATCTTAATGGTTAAACATTGAAGCGGAACAACTACTGAA 60
Db 1 ATGACTACTTACAAATTAATCTTAATGGTTAAACATTGAAGCGGAACAACTACTGAA 60
Qy 61 GCTGTTGATGCTGCTACTGCGAGCAAAAGTCTTCAACAATACGCTAACGCAACGGTGT 120
Db 61 GCTGTTGATGCTGCTACTGCTATTAAAGTCTTCAACAATACGCTAACGCAACGGTGT 120
Qy 121 GACGGTGAATGAGCTTACGACGATGCGACTAAGACCTTTACAGTTACTGAAACATCACCAT 180
Db 121 GACGGTGAATGAGCTTACGACGATGCGACTAAGACCTTTACAGTTACTGAAACATCACCAT 180
Qy 181 CATCACTAAGCTTGA 195
Db 181 CATCACTAAGCTTGA 195

RESULT 5
AAC86829
ID AAC86829 standard; DNA; 195 BP.
XX AC AAC86829;
XX DT 02-APR-2001 (first entry)
XX DE Nucleotide sequence of a mutated B1 domain of protein G.
XX KW B1 domain; protein G; Fab fragment; immunoglobulin G; IgG; Fc fragment;
XX KW knobs-into-holes binding site; ss.
XX OS Synthetic.
XX OS Streptococcus sp.
XX

```


ID XX AAC86822 standard; DNA; 195 BP.
AC XX AAC86822;
XX
DT 02-APR-2001 (first entry)
XX
DE Nucleotide sequence of a B1 domain of protein G with mutation K28A.
XX
XX B1 domain; protein G; Fab fragment; immunoglobulin G; IgG; Fc fragment;
KW knobs-into-holes binding site; ss.
XX
OS Synthetic.
OS Streptococcus sp.
XX
FH Key Location/Qualifiers
FT CDS 1..189
FT /*tag= a
FT /product= "B1 domain of protein G"
FT mat_peptide 4..189
FT /*tag= c
XX
PN WO200074728-A1.
XX
XX 14-DEC-2000.
XX
XX 02-JUN-2000; 2000WO-US015366.
XX
XX 04-JUN-1999; 99US-00326342.
XX
XX (UYDU-) UNIV DUKE.
XX
XX Hellingsa HW, Sloan DJ;
XX
XX WPI; 2001-045994/06.
XX P-PSDB; AAB30887.
XX
XX B1 domain polypeptide of bacterial protein G, useful for detection and
PT purification of antibody fragments, exhibits binding activity for Fab
PT fragment of an IgG but does not bind Fc fragment.
XX
XX Claim 34; Page 74; 83pp; English.
XX
XX The present sequence encodes a mutated B1 domain of the bacterial protein
CC G. The B1 domain exhibits binding activity for a Fab fragment of an
CC immunoglobulin (Ig) G but exhibits no binding activity for a Fc fragment
CC of an IgG. The B1 domain may further comprise a disrupted or mutated
CC "knobs-into-holes" binding site for a Fc fragment of an IgG. The Glu27
CC residue of the mature B1 domain acts as a charged knob on the surface of
CC the B1 domain, inserting a polar hole on the Fc fragment. A mutation of
CC this residue virtually abolishes stable complex formation. The B1 domain
CC can be immobilized to a solid phase support, and be used for purifying Fc
CC and Fab fragments of IgG from a warm-blooded vertebrate by affinity
CC chromatography. The B1 domain is also useful for detecting IgG or its
XX fragments in a clinical sample
SQ Sequence 195 BP; 67 A; 42 C; 38 G; 48 T; 0 U; 0 Other;
Query Match 95.9%; Score 187; DB 4; Length 195;
Best Local Similarity 97.4%; Pred. No. 1.9e-44;
Matches 190; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 ATGACTACTTACAAATTAACTCTTAATGGTAAACATTGAAAGCGGAAACAACACTACTGAA 60
DB 1 ATGACTACTTACAAATTAACTCTTAATGGTAAACATTGAAAGCGGAAACAACACTACTGAA 60
QY 61 GCTGTTGATGCTGCTACTGCGAGCGAAAGTCTTCAACAATAGCTTAACGACAAACGGTGT 120
DB 61 GCTGTTGATGCTGCTACTGCGAGCGGTCTTCAACAATAGCTTAACGACAAACGGTGT 120
QY 121 GACGGTGAATGACTTTACGACGATCGGACTAAGACCTTTACAGTTACTGAAACATCACCAT 180
DB 121 GACGGTGAATGACTTTACGACGATCGGACTAAGACCTTTACAGTTACTGAAACATCACCAT 180

QY 181 CATCACTAAGCTTGA 195
DB 181 CATCACTAAGCTTGA 195
RESULT 8
AAC86823
ID AAC86823 standard; DNA; 195 BP.
XX
XX AAC86823;
XX
XX 02-APR-2001 (first entry)
XX
XX Nucleotide sequence of a B1 domain of protein G with mutation K31A.
XX
XX B1 domain; protein G; Fab fragment; immunoglobulin G; IgG; Fc fragment;
KW knobs-into-holes binding site; ss.
XX
OS Synthetic.
OS Streptococcus sp.
XX
FH Key Location/Qualifiers
FT CDS 1..189
FT /*tag= a
FT /product= "B1 domain of protein G"
FT mat_peptide 4..189
FT /*tag= c
XX
PN WO200074728-A1.
XX
XX 14-DEC-2000.
XX
XX 02-JUN-2000; 2000WO-US015366.
XX
XX 04-JUN-1999; 99US-00326342.
XX
XX (UYDU-) UNIV DUKE.
XX
XX Hellingsa HW, Sloan DJ;
XX
XX WPI; 2001-045994/06.
XX P-PSDB; AAB30888.
XX
XX B1 domain polypeptide of bacterial protein G, useful for detection and
PT purification of antibody fragments, exhibits binding activity for Fab
PT fragment of an IgG but does not bind Fc fragment.
XX
XX Claim 34; Page 74-75; 83pp; English.
XX
XX The present sequence encodes a mutated B1 domain of the bacterial protein
CC G. The B1 domain exhibits binding activity for a Fab fragment of an
CC immunoglobulin (Ig) G but exhibits no binding activity for a Fc fragment
CC of an IgG. The B1 domain may further comprise a disrupted or mutated
CC "knobs-into-holes" binding site for a Fc fragment of an IgG. The Glu27
CC residue of the mature B1 domain acts as a charged knob on the surface of
CC the B1 domain, inserting a polar hole on the Fc fragment. A mutation of
CC this residue virtually abolishes stable complex formation. The B1 domain
CC can be immobilized to a solid phase support, and be used for purifying Fc
CC and Fab fragments of IgG from a warm-blooded vertebrate by affinity
CC chromatography. The B1 domain is also useful for detecting IgG or its
XX fragments in a clinical sample
SQ Sequence 195 BP; 67 A; 42 C; 38 G; 48 T; 0 U; 0 Other;
Query Match 95.9%; Score 187; DB 4; Length 195;
Best Local Similarity 97.4%; Pred. No. 1.9e-44;
Matches 190; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 ATGACTACTTACAAATTAACTCTTAATGGTAAACATTGAAAGCGGAAACAACACTACTGAA 60
DB 1 ATGACTACTTACAAATTAACTCTTAATGGTAAACATTGAAAGCGGAAACAACACTACTGAA 60
QY 61 GCTGTTGATGCTGCTACTGCGAGCGAAAGTCTTCAACAATAGCTTAACGACAAACGGTGT 120

```
Db 61 GCTGTTGATGCTGCTACTGCGAATAAAGTCTTCGCGCAATACGCTAACGCAACCGGTGTT 120
Qy 121 GACGGTGAATGGACTTACGACGATGCGACTTAAGACCTTTACAGTTACTGTAACATCACCAT 180
Db 121 GACGGTGAATGGACTTACGACGATGCGACTTAAGACCTTTACAGTTACTGTAACATCACCAT 180
Qy 181 CATCACTAAGCTTTGA 195
Db 181 CATCACTAAGCTTTGA 195

RESULT 9
AAC86824
ID AAC86824 standard; DNA; 195 BP.
AC AAC86824;
XX
XX 02-APR-2001 (first entry)
XX
XX Nucleotide sequence of a B1 domain of protein G with mutation N35A.
XX
XX B1 domain; protein G; Fab fragment; immunoglobulin G; IgG; Fc fragment;
XX knobs-into-holes binding site; ss.
XX
XX Synthetic.
XX Streptococcus sp.
XX
XX Key Location/Qualifiers
XX CDS 1..189
XX /tag= a
XX /product= "B1 domain of protein G"
XX mat_peptide 4..189
XX /tag= c
XX
XX WO200074728-A1.
XX
XX 14-DEC-2000.
XX
XX 02-JUN-2000; 2000WO-US015366.
XX
XX 04-JUN-1999; 99US-00326342.
XX (UYDU-) UNIV DUKE.
XX
XX Hellinga HW, Sloan DJ;
XX
XX WPI; 2001-049994/06.
XX P-PSDB; AAB30889.
XX
XX B1 domain polypeptide of bacterial protein G, useful for detection and
XX purification of antibody fragments, exhibits binding activity for Fab
XX fragment of an IgG but does not bind Fc fragment.
XX
XX Claim 34; Page 75; 83pp; English.
XX
XX The present sequence encodes a mutated B1 domain of the bacterial protein
XX G. The B1 domain exhibits binding activity for a Fab fragment of an
XX immunoglobulin (Ig) G but exhibits no binding activity for a Fc fragment
XX of an IgG. The B1 domain may further comprise a disrupted or mutated
XX "knobs-into-holes" binding site for a Fc fragment of an IgG. The Glu27
XX residue of the mature B1 domain acts as a charged knob on the surface of
XX the B1 domain, inserting a polar hole on the Fc fragment. A mutation of
XX this residue virtually abolishes stable complex formation. The B1 domain
XX can be immobilized to a solid phase support, and be used for purifying Fc
XX and Fab fragments of IgG from a warm-blooded vertebrate by affinity
XX chromatography. The B1 domain is also useful for detecting IgG or its
XX fragments in a clinical sample
XX
XX Sequence 195 BP; 68 A; 41 C; 38 G; 48 T; 0 U; 0 Other;
XX
XX Query Match 95.9%; Score 187; DB 4; Length 195;
XX Best Local Similarity 97.4%; Pred. No. 1.9e-44;
```

```
Matches 190; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 ATGACTACTTTACAAATTAATCCTTAATGGTAAACATTGAAAGCGCAACAACTACTGAA 60
Db 1 ATGACTACTTTACAAATTAATCCTTAATGGTAAACATTGAAAGCGCAACAACTACTGAA 60
Qy 61 GCTGTTGATGCTGCTACTGCGAATAAAGTCTTCGCGCAATACGCTAACGCAACCGGTGTT 120
Db 61 GCTGTTGATGCTGCTACTGCGAATAAAGTCTTCGCGCAATACGCTAACGCAACCGGTGTT 120
Qy 121 GACGGTGAATGGACTTACGACGATGCGACTTAAGACCTTTACAGTTACTGTAACATCACCAT 180
Db 121 GACGGTGAATGGACTTACGACGATGCGACTTAAGACCTTTACAGTTACTGTAACATCACCAT 180
Qy 181 CATCACTAAGCTTTGA 195
Db 181 CATCACTAAGCTTTGA 195

RESULT 10
AAC86826
ID AAC86826 standard; DNA; 195 BP.
XX
XX AAC86826;
XX
XX 02-APR-2001 (first entry)
XX
XX Nucleotide sequence of a B1 domain of protein G with mutation T44A/Y45A.
XX
XX B1 domain; protein G; Fab fragment; immunoglobulin G; IgG; Fc fragment;
XX knobs-into-holes binding site; ss.
XX
XX Synthetic.
XX Streptococcus sp.
XX
XX Key Location/Qualifiers
XX CDS 1..189
XX /tag= a
XX /product= "B1 domain of protein G"
XX mat_peptide 4..189
XX /tag= c
XX
XX WO200074728-A1.
XX
XX 14-DEC-2000.
XX
XX 02-JUN-2000; 2000WO-US015366.
XX
XX 04-JUN-1999; 99US-00326342.
XX (UYDU-) UNIV DUKE.
XX
XX Hellinga HW, Sloan DJ;
XX
XX WPI; 2001-049994/06.
XX P-PSDB; AAB30893.
XX
XX B1 domain polypeptide of bacterial protein G, useful for detection and
XX purification of antibody fragments, exhibits binding activity for Fab
XX fragment of an IgG but does not bind Fc fragment.
XX
XX Claim 34; Page 78; 83pp; English.
XX
XX The present sequence encodes a mutated B1 domain of the bacterial protein
XX G. The B1 domain exhibits binding activity for a Fab fragment of an
XX immunoglobulin (Ig) G but exhibits no binding activity for a Fc fragment
XX of an IgG. The B1 domain may further comprise a disrupted or mutated
XX "knobs-into-holes" binding site for a Fc fragment of an IgG. The Glu27
XX residue of the mature B1 domain acts as a charged knob on the surface of
XX the B1 domain, inserting a polar hole on the Fc fragment. A mutation of
XX this residue virtually abolishes stable complex formation. The B1 domain
XX can be immobilized to a solid phase support, and be used for purifying Fc
XX and Fab fragments of IgG from a warm-blooded vertebrate by affinity
```

CC chromatography. The B1 domain is also useful for detecting IgG or its
CC fragments in a clinical sample

XX Sequence 195 BP; 68 A; 41 C; 40 G; 46 T; 0 U; 0 Other;
SQ Query Match 94.3%; Score 183.8; DB 4; Length 195;
Best Local Similarity 96.4%; Pred. No. 1.6e-43;
Matches 188; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATGACTACTTCAAAATTAATCTTAATGGTTAAACATTGAAGCGGAAACAACTACTGAA 60
DB 1 ATGACTACTTCAAAATTAATCTTAATGGTTAAACATTGAAGCGGAAACAACTACTGAA 60
QY 61 GCTGTTGATGCTGCTACTGCGAGGAAAGTCTTCAAAACAATACGCTTAACGACAAACGGTGT 120
DB 61 GCTGTTGATGCTGCTACTGCGAGGAAAGTCTTCAAAACAATACGCTTAACGACAAACGGTGT 120
QY 121 GACGGTGAATGGACTTACGACGATCGGACTAAGACCTTTACAGTTACTGAAACATCACAT 180
DB 121 GACGGTGAATGGCGGCGGACGATCGGACTAAGACCTTTACAGTTACTGAAACATCACAT 180
QY 181 CATCACTAAGCTTGA 195
DB 181 CATCACTAAGCTTGA 195

RESULT 11
AAQ50453
ID AAQ50453 standard; DNA; 1308 BP.

AC AAQ50453;
XX 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 16-MAY-1994 (first entry)
XX

XX Sequence encoding immunoglobulin light chain binding protein.

XX Immunoglobulin; light chain; binding; identification; purification;
KW separation; ss.

XX Escherichia coli.

XX WO9322342-A1.

XX 11-NOV-1993.

XX 28-APR-1993; 93WO-SE000375.

XX 28-APR-1992; 92SE-00001331.

XX (HIGH-) HIGHTECH RECEPTOR AB.

XX Bjoerck L, Sjoebing U;

XX WPI; 1993-368722/46.

XX P-PSDB; AAR42994.

XX New protein L binding light chains of all immunoglobulin classes - for
PT binding purifying and identifying immunoglobulin, also related DNA,
PT vectors and host cells.

XX Claim 6; Page 39-40; 7lpp; English.

XX Protein L (AAR42993) is capable of binding to immunoglobulin G light
CC chains. It is useful for binding, separating (purifying) and identifying
CC immunoglobulin and for removing immunoglobulin molecules from serum. This
CC is the coding sequence of one hybrid protein of the L protein. The hybrid
CC proteins can bind all human immunoglobulin classes and many
CC immunoglobulins from other species. They are highly soluble and retain
CC their binding activity at high temperatures over a pH range of 3-10. They
CC can be immobilised without loss of activity. (Updated on 25-MAR-2003 to
CC correct PN field.) (Updated on 27-AUG-2003 to correct OS field.)

XX Sequence 1308 BP; 577 A; 204 C; 252 G; 275 T; 0 U; 0 Other;
SQ Query Match 84.0%; Score 163.8; DB 2; Length 1308;
Best Local Similarity 96.0%; Pred. No. 1.7e-37;
Matches 168; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 ATGACTACTTCAAAATTAATCTTAATGGTTAAACATTGAAGCGGAAACAACTACTGAA 60
DB 919 ATGACACTTTACAAATTAATCTTAATGGTTAAACATTGAAGCGGAAACAACTACTGAA 978
QY 61 GCTGTTGATGCTGCTACTGCGAGGAAAGTCTTCAAAACAATACGCTTAACGACAAACGGTGT 120
DB 979 GCTGTTGATGCTGCTACTGCGAGGAAAGTCTTCAAAACAATACGCTTAACGACAAACGGTGT 1038
QY 121 GACGGTGAATGGACTTACGACGATCGGACTAAGACCTTTACAGTTACTGAAACATC 175
DB 1039 GACGGTGAATGGACTTACGACGATCGGACTAAGACCTTTACAGTTACTGAAACATC 1093

RESULT 12
AAN94677
ID AAN94677 standard; DNA; 396 BP.

AC AAN94677;

XX 25-MAR-2003 (revised)
DT 04-JUL-1990 (first entry)
XX

XX Protein G gene variant.

XX Protein G; immunoglobulin; Fc receptor; ds.

XX Streptococcus sp.

XX WO8810306-A.

XX 29-DEC-1988.

XX 20-JUN-1988; 88WO-US002084.

XX 19-JUN-1987; 87US-00063959.

XX (GEMX) GENEX CORP.

XX (PHAA) PHARMACIA AB.

XX Farnestock SR;

XX WPI; 1989-023848/03.

XX P-PSDB; AAP94791.

XX Cloned protein G variant genes - expressing proteins having
PT immunoglobulin-binding properties of protein G and derived from
PT Streptococcus sp.

XX Claim 10; Page 86; 116pp; English.

XX Gene encodes protein G variant of non-pathogenic streptococcus allowing
CC isolation of the protein and variants, useful as bacterial Fc receptors
CC eg in purification and detection of Abs., screening of hybridoma clones
CC and treatment of disease. (Updated on 25-MAR-2003 to correct PA field.)
XX

XX Sequence 396 BP; 149 A; 87 C; 79 G; 81 T; 0 U; 0 Other;

Query Match 83.4%; Score 162.6; DB 1; Length 396;
Best Local Similarity 97.6%; Pred. No. 2.8e-37;
Matches 165; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 ACTTACAAATTAATCTTAATGGTTAAACATTGAAGCGGAAACAACTACTGAACTGTT 66

DB 31 ACTTACAAATTAATCTTAATGGTTAAACATTGAAGCGGAAACAACTACTGAACTGTT 90

QY 67 GATGCTGCTACTGCGAGGAAAGTCTTCAAAACAATACGCTTAACGACAAACGGTGTGACGGT 126

Db 91 GATGCTGTTACTGCGAAAAAGTCTTCAACAATACGCTAACGACAACGGTGTTCACGGT 150

Qy 127 GAATGGACTTACGAGTGGCGACTAAGACCTTTACAGTTACTGAACATC 175

Db 151 GAATGGACTTACGAGTGGCGACTAAGACCTTTACAGTTACTGAACAAAC 199

RESULT 13
AAQ06013
ID AAQ06013 standard: DNA: 396 BP.

RESULT 14

AC AAN91100;
XX
XX
DT 25-MAR-2003 (revised)
DT 04-JUL-1990 (first entry)
XX
XX DE Protein G gene variant.
XX
XX KW Protein G; immunoglobulin; Fc receptor; ds.
XX
XX OS Streptococcus sp.
XX
XX FN W08810306-A.
XX
XX PD 29-DEC-1988.
XX
XX PF 20-JUN-1988; 88WO-US002084.
XX
XX PR 19-JUN-1987; 87US-00063959.
XX
XX PA (GEMX) GENEX CORP.
XX PA (PHAA) PHARMACIA AB.
XX
XX FI Fahnestock SR;
XX
XX DR WPI; 1989-023848/03.
XX DR P-PSDB; AAP94784.
XX
XX PT Cloned protein G variant genes - expressing proteins having
XX PT immunoglobulin-binding properties of protein G and derived from
XX PT Streptococcus sp.
XX
XX PS Claim 3; Page 83; 116pp; English.
XX
XX CC Gene encodes protein G variant of non-pathogenic streptococcus allowing
XX CC isolation of the protein and variants, useful as bacterial Fc receptors
XX CC eg in purification and detection of Abs., screening of hybridoma clones
XX CC and treatment of disease. (Updated on 25-MAR-2003 to correct PA field.)
XX
XX SQ Sequence 435 BP; 155 A; 87 C; 90 G; 103 T; 0 U; 0 Other;

Query Match 83.4%; Score 162.6; DB 1; Length 435;
Best Local Similarity 97.6%; Pred. NO. 2.9e-37;
Matches 165; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 7 ACTTACAAATTAATCCTTAATGGTAAACATTTGAAGCGGAACAACTACTGAGCTGTT 66
Db 31 ACTTACAAATTAATCCTTAATGGTAAACATTTGAAGCGGAACAACTACTGAGCTGTT 90

Qy 67 GATGCTGCTACTGCAGCGAAAGTCTTCAAAACAATACGCTAACGACAAACGGTGTGACGGT 126
Db 91 GATGCTGCTACTGCAGAAAGTCTTCAAAACAATACGCTAACGACAAACGGTGTGACGGT 150

Qy 127 GAATGGACTTACGACGATGCGACTAAGACCTTTTACAGTTACTGACATC 175
Db 151 GAATGGACTTACGACGATGCGACTAAGACCTTTTACAGTTACTGACAAAC 199

Search completed: August 10, 2005, 22:28:39
Job time : 431 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2005, 15:15:58 ; Search time 1719 Seconds
(without alignments)

5496.665 Million cell updates/sec

Title: US-10-672-108-5

Perfect score: 195

Sequence: 1 atgactactacaataat.....accatcatcactaagcttga 195

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_btg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_ats.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	195	100.0	195	6	AR438294 Sequence
2	191.8	98.4	195	6	AR438293 Sequence
3	191.8	98.4	195	6	AR438300 Sequence
4	190.2	97.5	195	6	AR438301 Sequence
5	190.2	97.5	195	6	AR438302 Sequence
6	188.6	96.7	195	6	AR438298 Sequence
7	187	95.9	195	6	AR438295 Sequence
8	187	95.9	195	6	AR438296 Sequence
9	187	95.9	195	6	AR438297 Sequence
10	183.8	94.3	195	6	AR438299 Sequence
11	163.8	84.0	1308	6	AR079001 Sequence
12	163.8	84.0	1365	1	S50809 protein LG=
13	162.6	83.4	396	6	I09113 Sequence 10
14	162.6	83.4	435	6	I09106 Sequence 3
15	162.6	83.4	447	6	I09104 Sequence 1
16	162.6	83.4	495	6	I09112 Sequence 9
17	162.6	83.4	600	6	I09105 Sequence 2
18	162.6	83.4	606	6	I09111 Sequence 8
19	162.6	83.4	798	6	I09107 Sequence 4

20	162.6	83.4	810	6	I09109 Sequence 6
21	162.6	83.4	1469	1	SG1481GG
22	162.6	83.4	1469	6	A12446
23	162.6	83.4	1576	1	SG148PG
24	162.6	83.4	1950	1	STRSPGIGP
25	162.6	83.4	1950	6	I08536
26	162.6	83.4	1950	6	I09115
27	162.6	83.4	2384	1	SGSPG
28	162.6	83.4	2384	6	I08537
29	162.6	83.4	2456	1	SGPROTG
30	161.8	83.0	165	6	A12443
31	158	81.0	398	6	A69078
32	158	81.0	1422	6	I44708
33	158	81.0	1422	6	I52114
34	158	81.0	1683	6	I44709
35	158	81.0	1683	6	I52115
36	158	81.0	1695	6	I44711
37	158	81.0	1695	6	I52117
38	158	81.0	1722	6	I44712
39	158	81.0	1722	6	I52118
40	153.8	78.9	165	6	A12444
41	149	76.4	390	6	I09114
42	133.2	68.3	252	6	I09110
43	133.2	68.3	402	6	I09108
44	131	67.2	191	1	S62801
45	130	66.7	1555	1	STRMAG

ALIGNMENTS

RESULT 1
AR438294
LOCUS AR438294 195 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 5 from patent US 6663862.
ACCESSION AR438294
VERSION AR438294.1 GI:42662959
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 195)
AUTHORS Hellinga,H.W. and Sloan,D.J.
TITLE Reagents for detection and purification of antibody fragments
JOURNAL Patent: US 6663862-A 5 16-DEC-2003;
FEATURES
source
Location/Qualifiers
1..195
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 195; DB 6; Length 195;
Best Local Similarity 100.0%; Pred. No. 7.8e-41;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGACTACTTACAAATTAATCTTAAATGGTAAACATTTGAAAGCGGAAACAACTACTGAA 60
Db 1 ATGACTACTTACAAATTAATCTTAAATGGTAAACATTTGAAAGCGGAAACAACTACTGAA 60
Qy 61 GCTGTTGATGCTGCTACTGCGAGGAAAGTCTTCAACAATAGCTTAAGCTTAAGCAACGGTGT 120
Db 61 GCTGTTGATGCTGCTACTGCGAGGAAAGTCTTCAACAATAGCTTAAGCTTAAGCAACGGTGT 120
Qy 121 GACCGTGAATGAGCTTACGAGCGACTGAAGACCTTTACAGTTTACTGTAACATCACCAT 180
Db 121 GACCGTGAATGAGCTTACGAGCGACTGAAGACCTTTACAGTTTACTGTAACATCACCAT 180
Qy 181 CATCACTAAGCTTGA 195
Db 181 CATCACTAAGCTTGA 195
RESULT 2

```
AR438293
LOCUS AR438293 195 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 1 from patent US 6663862.
ACCESSION AR438293
VERSION AR438293.1 GI:42662958
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 195)
AUTHORS Hellings,H.W. and Sloan,D.J.
TITLE Reagents for detection and purification of antibody fragments
JOURNAL Patent: US 6663862-A 1 16-DEC-2003;
FEATURES
source
1..195
/mol_type="genomic DNA"
ORIGIN
Query Match 98.4%; Score 191.8; DB 6; Length 195;
Best Local Similarity 99.0%; Pred. No. 5.4e-40;
Matches 193; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGACTACTTACAAATTAATCCTTAATGGTAAACATTTGAAGCGGAAACAACTACTGAA 60
DB 1 ATGACTACTTACAAATTAATCCTTAATGGTAAACATTTGAAGCGGAAACAACTACTGAA 60
QY 61 GCTGTTGATGCTGCTACTGCGAGGAAAGTCTTCAAAACATTAACGCTAACGCAACGGTGTT 120
DB 61 GCTGTTGATGCTGCTACTGCGAGGAAAGTCTTCAAAACATTAACGCTAACGCAACGGTGTT 120
QY 121 GACGGTGAATGGACTTACGACGATGCGACTAAGACCTTTTACAGTTACTGAACATCACCAT 180
DB 121 GACGGTGAATGGACTTACGACGATGCGACTAAGACCTTTTACAGTTACTGAACATCACCAT 180
QY 181 CATCACTAAGCTTGA 195
DB 181 CATCACTAAGCTTGA 195
RESULT 3
AR438300
LOCUS AR438300 195 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 19 from patent US 6663862.
ACCESSION AR438300
VERSION AR438300.1 GI:42662965
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 195)
AUTHORS Hellings,H.W. and Sloan,D.J.
TITLE Reagents for detection and purification of antibody fragments
JOURNAL Patent: US 6663862-A 19 16-DEC-2003;
FEATURES
source
1..195
/mol_type="genomic DNA"
ORIGIN
Query Match 98.4%; Score 191.8; DB 6; Length 195;
Best Local Similarity 99.0%; Pred. No. 5.4e-40;
Matches 193; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGACTACTTACAAATTAATCCTTAATGGTAAACATTTGAAGCGGAAACAACTACTGAA 60
DB 1 ATGACTACTTACAAATTAATCCTTAATGGTAAACATTTGAAGCGGAAACAACTACTGAA 60
QY 61 GCTGTTGATGCTGCTACTGCGAGGAAAGTCTTCAAAACATTAACGCTAACGCAACGGTGTT 120
DB 61 GCTGTTGATGCTGCTACTGCGAGGAAAGTCTTCAAAACATTAACGCTAACGCAACGGTGTT 120
QY 121 GACGGTGAATGGACTTACGACGATGCGACTAAGACCTTTTACAGTTACTGAACATCACCAT 180
DB 121 GACGGTGAATGGACTTACGACGATGCGACTAAGACCTTTTACAGTTACTGAACATCACCAT 180
QY 181 CATCACTAAGCTTGA 195
DB 181 CATCACTAAGCTTGA 195
AR438293
LOCUS AR438293 195 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 1 from patent US 6663862.
ACCESSION AR438293
VERSION AR438293.1 GI:42662958
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 195)
AUTHORS Hellings,H.W. and Sloan,D.J.
TITLE Reagents for detection and purification of antibody fragments
JOURNAL Patent: US 6663862-A 1 16-DEC-2003;
FEATURES
source
1..195
/mol_type="genomic DNA"
ORIGIN
Query Match 98.4%; Score 191.8; DB 6; Length 195;
Best Local Similarity 99.0%; Pred. No. 5.4e-40;
Matches 193; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGACTACTTACAAATTAATCCTTAATGGTAAACATTTGAAGCGGAAACAACTACTGAA 60
DB 1 ATGACTACTTACAAATTAATCCTTAATGGTAAACATTTGAAGCGGAAACAACTACTGAA 60
QY 61 GCTGTTGATGCTGCTACTGCGAGGAAAGTCTTCAAAACATTAACGCTAACGCAACGGTGTT 120
DB 61 GCTGTTGATGCTGCTACTGCGAGGAAAGTCTTCAAAACATTAACGCTAACGCAACGGTGTT 120
QY 121 GACGGTGAATGGACTTACGACGATGCGACTAAGACCTTTTACAGTTACTGAACATCACCAT 180
DB 121 GACGGTGAATGGACTTACGACGATGCGACTAAGACCTTTTACAGTTACTGAACATCACCAT 180
QY 181 CATCACTAAGCTTGA 195
DB 181 CATCACTAAGCTTGA 195
AR438300
LOCUS AR438300 195 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 23 from patent US 6663862.
ACCESSION AR438300
VERSION AR438302.1 GI:42662967
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 195)
AUTHORS Hellings,H.W. and Sloan,D.J.
TITLE Reagents for detection and purification of antibody fragments
JOURNAL Patent: US 6663862-A 23 16-DEC-2003;
FEATURES
source
1..195
/mol_type="genomic DNA"
ORIGIN
Query Match 97.5%; Score 190.2; DB 6; Length 195;
Best Local Similarity 98.5%; Pred. No. 1.4e-39;
Matches 192; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ATGACTACTTACAAATTAATCCTTAATGGTAAACATTTGAAGCGGAAACAACTACTGAA 60
DB 1 ATGACTACTTACAAATTAATCCTTAATGGTAAACATTTGAAGCGGAAACAACTACTGAA 60
QY 61 GCTGTTGATGCTGCTACTGCGAGGAAAGTCTTCAAAACATTAACGCTAACGCAACGGTGTT 120
DB 61 GCTGTTGATGCTGCTACTGCGATTTAAAGTCTTCAAAACATTAACGCTAACGCAACGGTGTT 120
QY 121 GACGGTGAATGGACTTACGACGATGCGACTAAGACCTTTTACAGTTACTGAACATCACCAT 180
DB 121 GACGGTGAATGGACTTACGACGATGCGACTAAGACCTTTTACAGTTACTGAACATCACCAT 180
QY 181 CATCACTAAGCTTGA 195
DB 181 CATCACTAAGCTTGA 195
AR438302
LOCUS AR438302 195 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 23 from patent US 6663862.
ACCESSION AR438302
VERSION AR438302.1 GI:42662967
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 195)
AUTHORS Hellings,H.W. and Sloan,D.J.
TITLE Reagents for detection and purification of antibody fragments
JOURNAL Patent: US 6663862-A 23 16-DEC-2003;
FEATURES
source
1..195
/mol_type="genomic DNA"
ORIGIN
Query Match 97.5%; Score 190.2; DB 6; Length 195;
Best Local Similarity 98.5%; Pred. No. 1.4e-39;
Matches 192; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

Qy 1 ATGACTACTTACAAATTAATCCTTAATGGTAAACATTGAAAGCGGAAACAACTACTGAA 60
Db 1 ATGACTACTTACAAATTAATCCTTAATGGTAAACATTGAAAGCGGAAACAACTACTGAA 60
Qy 61 GCTGTTGATGCTGCTACTGCGAGCGAAAGTCTTCAACAATACGCTTAACGACAACGGTGT 120
Db 61 GCTGTTGATGCTGCTACTGCGAAATTAATGCTTCAACAATACGCTTAACGACAACGGTGT 120
Qy 121 GACGCTGAATGCGACTTACGAGCGATGCGACTTAAGACCTTTTACAGTTACTGAAACATCACC 180
Db 121 GACGCTGAATGCGACTTACGAGCGATGCGACTTAAGACCTTTTACAGTTACTGAAACATCACC 180
Qy 181 CATCACTAAGCTTGA 195
Db 181 CATCACTAAGCTTGA 195
RESULT 6
AR438298 LOCUS 195 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 15 from patent US 6663862.
ACCESSION AR438298
VERSION AR438298.1 GI:42662963
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 195)
AUTHORS Hellinga,H.W. and Sloan,D.J.
TITLE Reagents for detection and purification of antibody fragments
JOURNAL Patent: US 6663862-A 15 16-DEC-2003;
FEATURES Location/Qualifiers
source 1..195
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 96.7%; Score 188.6; DB 6; Length 195;
Best Local Similarity 97.9%; Pred. No. 3.7e-39;
Matches 191; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 ATGACTACTTACAAATTAATCCTTAATGGTAAACATTGAAAGCGGAAACAACTACTGAA 60
Db 1 ATGACTACTTACAAATTAATCCTTAATGGTAAACATTGAAAGCGGAAACAACTACTGAA 60
Qy 61 GCTGTTGATGCTGCTACTGCGAGCGAAAGTCTTCAACAATACGCTTAACGACAACGGTGT 120
Db 61 GCTGTTGATGCTGCTACTGCGAGCGAAATTAATGCTTCAACAATACGCTTAACGACAACGGTGT 120
Qy 121 GACGCTGAATGCGACTTACGAGCGATGCGACTTAAGACCTTTTACAGTTACTGAAACATCACC 180
Db 121 GACGCTGAAGCGACTTACGAGCGATGCGACTTAAGACCTTTTACAGTTACTGAAACATCACC 180
Qy 181 CATCACTAAGCTTGA 195
Db 181 CATCACTAAGCTTGA 195
RESULT 7
AR438295 LOCUS 195 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 7 from patent US 6663862.
ACCESSION AR438295
VERSION AR438295.1 GI:42662960
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 195)
AUTHORS Hellinga,H.W. and Sloan,D.J.
TITLE Reagents for detection and purification of antibody fragments
JOURNAL Patent: US 6663862-A 7 16-DEC-2003;
FEATURES Location/Qualifiers

source 1..195
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 95.9%; Score 187; DB 6; Length 195;
Best Local Similarity 97.4%; Pred. No. 9.9e-39;
Matches 190; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 ATGACTACTTACAAATTAATCCTTAATGGTAAACATTGAAAGCGGAAACAACTACTGAA 60
Db 1 ATGACTACTTACAAATTAATCCTTAATGGTAAACATTGAAAGCGGAAACAACTACTGAA 60
Qy 61 GCTGTTGATGCTGCTACTGCGAGCGAAAGTCTTCAACAATACGCTTAACGACAACGGTGT 120
Db 61 GCTGTTGATGCTGCTACTGCGAGAAAGCGGTCTTCAACAATACGCTTAACGACAACGGTGT 120
Qy 121 GACGCTGAATGCGACTTACGAGCGATGCGACTTAAGACCTTTTACAGTTACTGAAACATCACC 180
Db 121 GACGCTGAATGCGACTTACGAGCGATGCGACTTAAGACCTTTTACAGTTACTGAAACATCACC 180
Qy 181 CATCACTAAGCTTGA 195
Db 181 CATCACTAAGCTTGA 195
RESULT 8
AR438296 LOCUS 195 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 9 from patent US 6663862.
ACCESSION AR438296
VERSION AR438296.1 GI:42662961
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 195)
AUTHORS Hellinga,H.W. and Sloan,D.J.
TITLE Reagents for detection and purification of antibody fragments
JOURNAL Patent: US 6663862-A 9 16-DEC-2003;
FEATURES Location/Qualifiers
source 1..195
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 95.9%; Score 187; DB 6; Length 195;
Best Local Similarity 97.4%; Pred. No. 9.9e-39;
Matches 190; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 ATGACTACTTACAAATTAATCCTTAATGGTAAACATTGAAAGCGGAAACAACTACTGAA 60
Db 1 ATGACTACTTACAAATTAATCCTTAATGGTAAACATTGAAAGCGGAAACAACTACTGAA 60
Qy 61 GCTGTTGATGCTGCTACTGCGAGCGAAAGTCTTCAACAATACGCTTAACGACAACGGTGT 120
Db 61 GCTGTTGATGCTGCTACTGCGAGAAAGTCTTTCGCGCAATACGCTTAACGACAACGGTGT 120
Qy 121 GACGCTGAATGCGACTTACGAGCGATGCGACTTAAGACCTTTTACAGTTACTGAAACATCACC 180
Db 121 GACGCTGAATGCGACTTACGAGCGATGCGACTTAAGACCTTTTACAGTTACTGAAACATCACC 180
Qy 181 CATCACTAAGCTTGA 195
Db 181 CATCACTAAGCTTGA 195
RESULT 9
AR438297 LOCUS 195 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 11 from patent US 6663862.
ACCESSION AR438297
VERSION AR438297.1 GI:42662962

```
KEYWORDS      Unknown.
SOURCE        Unknown.
ORGANISM      Unclassified.
REFERENCE     1 (bases 1 to 195)
AUTHORS       Hellings,H.W. and Sloan,D.J.
TITLE         Reagents for detection and purification of antibody fragments
JOURNAL      Patent: US 6663862-A 11 16-DEC-2003;
FEATURES     Location/Qualifiers
              1..195
              /organism="unknown"
              /mol_type="genomic DNA"

ORIGIN
Query Match      95.9%; Score 187; DB 6; Length 195;
Best Local Similarity 97.4%; Pred. No. 9.9e-39;
Matches 190; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGACTACTTACAAATTAATCCTTAATGGTAAACATTGAAAGCGGAAACAACTACTGAA 60
    |||||
Db 1 ATGACTACTTACAAATTAATCCTTAATGGTAAACATTGAAAGCGGAAACAACTACTGAA 60

QY 61 GCTGTTGATGCTGCTACTGAGCGAAAGTCTTCAACAATACGCTTAACGACACGGTGT 120
    |||||
Db 61 GCTGTTGATGCTGCTACTGAGCGAAAGTCTTCAACAATACGCTGCGGACACGGTGT 120

QY 121 GACGGTGAATGGACTTACGACGATGCGACTTAAGACCTTTTACAGTTACTGAACATCACCAT 180
    |||||
Db 121 GACGGTGAATGGACTTACGACGATGCGACTTAAGACCTTTTACAGTTACTGAACATCACCAT 180

QY 181 CATCACTAAGCTTGA 195
    |||||
Db 181 CATCACTAAGCTTGA 195

RESULT 10
AR438299
LOCUS      AR438299      195 bp      DNA      linear      PAT 20-FEB-2004
DEFINITION Sequence 17 from patent US 6663862.
ACCESSION AR438299
VERSION   AR438299.1 GI:42662964
KEYWORDS
SOURCE    Unknown.
ORGANISM  Unclassified.
REFERENCE 1 (bases 1 to 195)
AUTHORS   Hellings,H.W. and Sloan,D.J.
TITLE     Reagents for detection and purification of antibody fragments
JOURNAL   Patent: US 6663862-A 17 16-DEC-2003;
FEATURES  Location/Qualifiers
          1..195
          /organism="unknown"
          /mol_type="genomic DNA"

ORIGIN
Query Match      94.3%; Score 183.8; DB 6; Length 195;
Best Local Similarity 96.4%; Pred. No. 6.8e-38;
Matches 188; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATGACTACTTACAAATTAATCCTTAATGGTAAACATTGAAAGCGGAAACAACTACTGAA 60
    |||||
Db 1 ATGACTACTTACAAATTAATCCTTAATGGTAAACATTGAAAGCGGAAACAACTACTGAA 60

QY 61 GCTGTTGATGCTGCTACTGAGCGAAAGTCTTCAACAATACGCTTAACGACACGGTGT 120
    |||||
Db 61 GCTGTTGATGCTGCTACTGAGCGAAAGTCTTCAACAATACGCTTAACGACACGGTGT 120

QY 121 GACGGTGAATGGACTTACGACGATGCGACTTAAGACCTTTTACAGTTACTGAACATCACCAT 180
    |||||
Db 121 GACGGTGAATGGCGGCGGACGATGCGACTTAAGACCTTTTACAGTTACTGAACATCACCAT 180

QY 181 CATCACTAAGCTTGA 195
    |||||

KEYWORDS      Unknown.
SOURCE        Unknown.
ORGANISM      Unclassified.
REFERENCE     1 (bases 1 to 195)
AUTHORS       Hellings,H.W. and Sloan,D.J.
TITLE         Reagents for detection and purification of antibody fragments
JOURNAL      Patent: US 6663862-A 11 16-DEC-2003;
FEATURES     Location/Qualifiers
              1..195
              /organism="unknown"
              /mol_type="genomic DNA"

ORIGIN
Query Match      95.9%; Score 187; DB 6; Length 195;
Best Local Similarity 97.4%; Pred. No. 9.9e-39;
Matches 190; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGACTACTTACAAATTAATCCTTAATGGTAAACATTGAAAGCGGAAACAACTACTGAA 60
    |||||
Db 1 ATGACTACTTACAAATTAATCCTTAATGGTAAACATTGAAAGCGGAAACAACTACTGAA 60

QY 61 GCTGTTGATGCTGCTACTGAGCGAAAGTCTTCAACAATACGCTTAACGACACGGTGT 120
    |||||
Db 61 GCTGTTGATGCTGCTACTGAGCGAAAGTCTTCAACAATACGCTGCGGACACGGTGT 120

QY 121 GACGGTGAATGGACTTACGACGATGCGACTTAAGACCTTTTACAGTTACTGAACATCACCAT 180
    |||||
Db 121 GACGGTGAATGGACTTACGACGATGCGACTTAAGACCTTTTACAGTTACTGAACATCACCAT 180

QY 181 CATCACTAAGCTTGA 195
    |||||

KEYWORDS      Unknown.
SOURCE        Unknown.
ORGANISM      Unclassified.
REFERENCE     1 (bases 1 to 1308)
AUTHORS       Bjorck,L. and Sjobring,U.
TITLE         Protein L and hybrid proteins thereof
JOURNAL      Patent: US 5965390-A 4 12-OCT-1999;
FEATURES     Location/Qualifiers
              1..1308
              /organism="unknown"
              /mol_type="unassigned DNA"

ORIGIN
Query Match      84.0%; Score 163.8; DB 6; Length 1308;
Best Local Similarity 96.0%; Pred. No. 1.3e-32;
Matches 168; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATGACTACTTACAAATTAATCCTTAATGGTAAACATTGAAAGCGGAAACAACTACTGAA 60
    |||||
Db 919 ATGGACACTTACAAATTAATCCTTAATGGTAAACATTGAAAGCGGAAACAACTACTGAA 978

QY 61 GCTGTTGATGCTGCTACTGAGCGAAAGTCTTCAACAATACGCTTAACGACACGGTGT 120
    |||||
Db 979 GCTGTTGATGCTGCTACTGAGCGAAAGTCTTCAACAATACGCTTAACGACACGGTGT 1038

QY 121 GACGGTGAATGGACTTACGACGATGCGACTTAAGACCTTTTACAGTTACTGAACATC 175
    |||||
Db 1039 GACGGTGAATGGACTTACGACGATGCGACTTAAGACCTTTTACAGTTACTGAACATC 1093

RESULT 12
S50809
LOCUS      S50809      1365 bp      mRNA      linear      BCT 02-MAR-2001
DEFINITION protein Ig-immunoglobulin binding protein {immunoglobulin binding
ACCESSION S50809
VERSION   S50809.1 GI:261705
KEYWORDS  Finnegoldia magna
SOURCE    Finnegoldia magna
ORGANISM  Bacteria; Firmicutes; Clostridia; Clostridiales;
          Peptostreptococaceae; Finnegoldia.
REFERENCE 1 (bases 1 to 1365)
AUTHORS   Kihlberg,B.M., Sjobring,U., Kastern,W. and Bjorck,L.
TITLE     Protein Ig: a hybrid molecule with unique immunoglobulin binding
          properties
JOURNAL   J. Biol. Chem. 267 (35), 25583-25588 (1992)
MEDLINE   93094283
PUBMED    1460053
REMARK    GenBank staff at the National Library of Medicine created this
          entry [NCBI gibbsq 120302] from the original journal article.
FEATURES  Location/Qualifiers
          1..987
          /organism="Finnegoldia magna"
          /mol_type="mRNA"
          /db_xref="taxon:1260"
          /focus
          988..1365
          /organism="Streptococcus sp."
          /mol_type="mRNA"
          /db_xref="taxon:1306"
          1..1365
          /note="recombinant immunoglobulin binding protein; hybrid

CDS

```

construct shows no termination codon"

```
/codon_start=1
/transl_table=11
/product="protein LG"
/protein_id="AAA03280.1"
/db_xref="GI:261706"
/translation="MKKTAIAIAVAFATVAQAQVENKEETPETPTDSBEVITIK
ANLIFPANGSTQATFKGTGFEKATSEAYAYADTLKONGEYTVDVADKGYTLNIFKAG
EKTPEEKEVTIKANLIYADKGTOTAEFKGTPEATAEAVRYADALDKONGEYTVDV
ADKGYTLNIFKAGKEPTPEEKEVTIKANLIYADKGTOTAEFKGTPEATAEAVRYA
DLAKENGKTVADVADKGYTLNIFKAGKEPTPEEKEVTIKANLIYADKGTOTAEFK
GTFAEATAEAVRYADALDKENGKTVADLEGGYTLNIFKAGKVDKPEPMDYTKLI
LNGTKLGEETTTAEVDAATAEKVKQYANDNGVDGEYTDATKTFVTVEKPEVIDAS
ELTPAVTYYKLIVNGKTLKGTTTKAVDAETAEKAPKQYANDNGVDGVWYDDATKTF
TVTEM"
```

ORIGIN

Query Match 84.0%; Score 163.8; DB 1; Length 1365;
Best Local Similarity 96.0%; Pred. No. 1.3e-32;
Matches 168; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ATGACTACTTACAAATTAATCCTTAATGGTAAACATTTGAAGCGGAAACAACACTACTGAA 60
Db 982 ATGACACTTACAAATTAATCCTTAATGGTAAACATTTGAAGCGGAAACAACACTACTGAA 1041
|||
Qy 61 GCTGTTGATGCTGCTACTGCGAGCAAGTCTTCAAAACAATACGCTAACGCAACCGGTGTT 120
Db 1042 GCTGTTGATGCTGCTACTGCGAGCAAGTCTTCAAAACAATACGCTAACGCAACCGGTGTT 1101
|||
Qy 121 GACGGTGAATGGACTTACGACGATCGGACTAAGACCTTTACAGTTACTGTAACATC 175
Db 1102 GACGGTGAATGGACTTACGACGATCGGACTAAGACCTTTACAGTTACTGTAACAAAC 1156
|||

RESULT 13

I09113 LOCUS 396 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 10 from Patent WO 8810306.
ACCESSION I09113
VERSION I09113.1 GI:588185
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 396)
AUTHORS Fahnestock, S.R.
TITLE CLONED PROTEIN G VARIANT GENES AND THE PROTEIN G VARIANTS EXPRESSED THEREFROM
JOURNAL Patent: WO 8810306-A 10 29-DEC-1988;
FEATURES Location/Qualifiers
source 1..396
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 83.4%; Score 162.6; DB 6; Length 396;
Best Local Similarity 97.6%; Pred. No. 2.6e-32;
Matches 165; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 7 ACTTACAAATTAATCCTTAATGGTAAACATTTGAAGCGGAAACAACACTACTGAAAGCTGTT 66
Db 31 ACTTACAAATTAATCCTTAATGGTAAACATTTGAAGCGGAAACAACACTACTGAAAGCTGTT 90
|||
Qy 67 GATGCTGCTACTGCGAGCAAGTCTTCAAAACAATACGCTAACGCAACCGGTGTTGACCGT 126
Db 91 GATGCTGCTACTGCGAGCAAGTCTTCAAAACAATACGCTAACGCAACCGGTGTTGACCGT 150
|||
Qy 127 GAATGGACTTACGACGATGCGACTAAGACCTTTACAGTTACTGTAACATC 175
Db 151 GAATGGACTTACGACGATGCGACTAAGACCTTTACAGTTACTGTAACAAAC 199
|||

RESULT 14

I09106

LOCUS 435 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 3 from Patent WO 8810306.
ACCESSION I09106
VERSION I09106.1 GI:588178
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 435)
AUTHORS Fahnestock, S.R.
TITLE CLONED PROTEIN G VARIANT GENES AND THE PROTEIN G VARIANTS EXPRESSED THEREFROM
JOURNAL Patent: WO 8810306-A 3 29-DEC-1988;
FEATURES Location/Qualifiers
source 1..435
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 83.4%; Score 162.6; DB 6; Length 435;
Best Local Similarity 97.6%; Pred. No. 2.6e-32;
Matches 165; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 7 ACTTACAAATTAATCCTTAATGGTAAACATTTGAAGCGGAAACAACACTACTGAAAGCTGTT 66
Db 31 ACTTACAAATTAATCCTTAATGGTAAACATTTGAAGCGGAAACAACACTACTGAAAGCTGTT 90
|||
Qy 67 GATGCTGCTACTGCGAGCAAGTCTTCAAAACAATACGCTAACGCAACCGGTGTTGACCGT 126
Db 91 GATGCTGCTACTGCGAGCAAGTCTTCAAAACAATACGCTAACGCAACCGGTGTTGACCGT 150
|||
Qy 127 GAATGGACTTACGACGATGCGACTAAGACCTTTACAGTTACTGTAACATC 175
Db 151 GAATGGACTTACGACGATGCGACTAAGACCTTTACAGTTACTGTAACAAAC 199
|||

RESULT 15

I09104 LOCUS 447 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 1 from Patent WO 8810306.
ACCESSION I09104
VERSION I09104.1 GI:588176
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 447)
AUTHORS Fahnestock, S.R.
TITLE CLONED PROTEIN G VARIANT GENES AND THE PROTEIN G VARIANTS EXPRESSED THEREFROM
JOURNAL Patent: WO 8810306-A 1 29-DEC-1988;
FEATURES Location/Qualifiers
source 1..447
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 83.4%; Score 162.6; DB 6; Length 447;
Best Local Similarity 97.6%; Pred. No. 2.6e-32;
Matches 165; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 7 ACTTACAAATTAATCCTTAATGGTAAACATTTGAAGCGGAAACAACACTACTGAAAGCTGTT 66
Db 28 ACTTACAAATTAATCCTTAATGGTAAACATTTGAAGCGGAAACAACACTACTGAAAGCTGTT 87
|||
Qy 67 GATGCTGCTACTGCGAGCAAGTCTTCAAAACAATACGCTAACGCAACCGGTGTTGACCGT 126
Db 88 GATGCTGCTACTGCGAGCAAGTCTTCAAAACAATACGCTAACGCAACCGGTGTTGACCGT 147
|||
Qy 127 GAATGGACTTACGACGATGCGACTAAGACCTTTACAGTTACTGTAACATC 175
Db 148 GAATGGACTTACGACGATGCGACTAAGACCTTTACAGTTACTGTAACAAAC 196
|||

Search completed: August 10, 2005, 22:57:30
Job time : 1720 secs
